



DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITE DE COOPERATION EN MATIÈRE DE BREVETS (PCT)

(51) Classification internationale des brevets ⁶ : C12N 15/48, C12Q 1/70, C07K 14/15, A61K 31/70	A1	(11) Numéro de publication internationale: WO 99/67395 (43) Date de publication internationale: 29 décembre 1999 (29.12.99)
(21) Numéro de la demande internationale: PCT/FR99/01513 (22) Date de dépôt international: 23 juin 1999 (23.06.99) (30) Données relatives à la priorité: 98/07920 23 juin 1998 (23.06.98) FR (71) Déposant (pour tous les Etats désignés sauf US): INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE-INSERM [FR/FR]; 101, rue de Tolbiac, F-75654 Paris Cedex 13 (FR). (72) Inventeurs; et (75) Inventeurs/Déposants (US seulement): ALLIEL, Patrick, M. [FR/FR]; 4, rue Lazare Carnot, F-92140 Clamart (FR). PERIN, Jean-Pierre [FR/FR]; 182, rue d'Aulnay, F-92350 Le Plessis-Robinson (FR). RIEGER, François [FR/FR]; 38 bis, boulevard de la République, F-92100 Boulogne (FR). (74) Mandataire: CABINET ORES; 6, avenue de Messine, F-75008 Paris (FR).		(81) Etats désignés: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, brevet ARIPO (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), brevet eurasién (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Publiée <i>Avec rapport de recherche internationale.</i> <i>Avant l'expiration du délai prévu pour la modification des</i> <i>revendications, sera republiée si des modifications sont</i> <i>reçues.</i>
(54) Title: NUCLEIC SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HUMAN ENDOGENOUS RETROVIRAL MOTIFS, AND THEIR USES		
(54) Titre: FAMILLE DE SEQUENCES NUCLEIQUES ET DE SEQUENCES PROTEIQUES DEDUITES PRESENTANT DES MOTIFS RETROVIRAUX ENDOGENES HUMAINS, ET LEURS APPLICATIONS		
(57) Abstract <p>The invention concerns a novel nucleic sequence and deduced protein sequence family with whole or partial human endogenous retroviral motifs. The invention also concerns the detection and/or the use of said nucleic sequences and said corresponding protein sequences or fragments of said sequences, for diagnostic, prophylactic and therapeutic uses, in particular for neuropathological conditions with autoimmune constituent such as multiple sclerosis. Said purified nucleic acid sequences comprise all or part of a sequence coding for a human endogenous retroviral sequence having at least <i>env</i>-type retroviral motifs, corresponding to the sequence SEQ ID NO:1 or to a sequence having a homology level with said sequence SEQ ID NO:1 not less than 80 % on more than 190 nucleotides or not less than 70 % on more than 600 nucleotides for <i>env</i>-type domains. The invention further concerns the use of the flanking or adjacent sequences of said sequences and controlled by the latter, as diagnostic reagents.</p>		
(57) Abrégé <p>Nouvelle famille de séquences nucléiques et de séquences protéiques déduites, qui présentent des motifs rétroviraux endogènes humains complets ou partiels. Détection et/ou utilisation desdites séquences nucléiques et desdites séquences protéiques correspondantes ou de fragments de ces séquences, dans le cadre d'applications diagnostiques, prophylactiques et thérapeutiques, en particulier pour des neuropathologies à composante auto-immune comme la sclérose en plaques. Lesdites séquences d'acide nucléique purifiées comprennent tout ou partie d'une séquence codant pour une séquence rétrovirale endogène humaine, qui présente au moins des motifs rétroviraux de type <i>env</i>, répondant à la séquence SEQ ID NO:1 ou à une séquence présentant un niveau d'homologie avec ladite séquence SEQ ID NO:1 supérieur ou égal à 80 % sur plus de 190 nucléotides ou supérieur ou égal à 70 % sur plus de 600 nucléotides pour les domaines de type <i>env</i>. Utilisation des séquences flanquantes ou adjacentes desdites séquences et contrôlées par ces dernières, comme réactifs de diagnostic.</p>		

UNIQUEMENT A TITRE D'INFORMATION

Codes utilisés pour identifier les Etats parties au PCT, sur les pages de couverture des brochures publiant des demandes internationales en vertu du PCT.

AL	Albanie	ES	Espagne	LS	Lesotho	SI	Slovénie
AM	Arménie	FI	Finlande	LT	Lituanie	SK	Slovaquie
AT	Autriche	FR	France	LU	Luxembourg	SN	Sénégal
AU	Australie	GA	Gabon	LV	Lettonie	SZ	Swaziland
AZ	Azerbaïdjan	GB	Royaume-Uni	MC	Monaco	TD	Tchad
BA	Bosnie-Herzégovine	GE	Géorgie	MD	République de Moldova	TG	Togo
BB	Barbade	GH	Ghana	MG	Madagascar	TJ	Tadjikistan
BE	Belgique	GN	Guinée	MK	Ex-République yougoslave de Macédoine	TM	Turkménistan
BF	Burkina Faso	GR	Grèce	ML	Mali	TR	Turquie
BG	Bulgarie	HU	Hongrie	MN	Mongolie	TT	Trinité-et-Tobago
BJ	Bénin	IE	Irlande	MR	Mauritanie	UA	Ukraine
BR	Brésil	IL	Israël	MW	Malawi	UG	Ouganda
BY	Bélarus	IS	Islande	MX	Mexique	US	Etats-Unis d'Amérique
CA	Canada	IT	Italie	NE	Niger	UZ	Ouzbékistan
CF	République centrafricaine	JP	Japon	NL	Pays-Bas	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norvège	YU	Yougoslavie
CH	Suisse	KG	Kirghizistan	NZ	Nouvelle-Zélande	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	République populaire démocratique de Corée	PL	Pologne		
CM	Cameroun	KR	République de Corée	PT	Portugal		
CN	Chine	KZ	Kazakstan	RO	Roumanie		
CU	Cuba	LC	Sainte-Lucie	RU	Fédération de Russie		
CZ	République tchèque	LI	Liechtenstein	SD	Soudan		
DE	Allemagne	LK	Sri Lanka	SE	Suède		
DK	Danemark	LR	Libéria	SG	Singapour		
EE	Estonie						

FAMILLE DE SEQUENCES NUCLEIQUES ET DE SEQUENCES PROTEIQUES DEDUITES PRESENTANT DES MOTIFS RETROVIRAUX ENDOGENES HUMAINS, ET LEURS APPLICATIONS

5 La présente invention est relative à une nouvelle famille de séquences nucléiques et de séquences protéiques déduites, qui présentent des motifs rétroviraux endogènes humains complets ou partiels, ainsi que des séquences flanquantes ou adjacentes desdites séquences, et contrôlées par ces dernières : modification de l'expression ou altération de la structure (polyadénylation, épissage alternatif...) desdites séquences flanquantes.

 L'invention est également relative à la détection et/ou à l'utilisation desdites séquences nucléiques et desdites séquences protéiques correspondantes, dans le cadre d'applications diagnostiques, prophylactiques et thérapeutiques, en particulier pour des neuropathologies à composante autoimmune comme la sclérose en plaques.

15 L'invention concerne aussi l'obtention de sondes nucléiques double brins et simple brin anti-sens, de ribozymes, aptes à moduler la réplication virale (T.R. Cech, *Science*, 1987, **236**, 1532-1539 ; R.H. Symons, *Trends Biochem. Sci.*, 1989, **14**, 445-450) des molécules recombinantes correspondantes, et des anticorps associés.

 Les rétrovirus sont des virus qui se répliquent uniquement en utilisant la voie inverse du traitement classique de l'information génétique. Ce processus, nommé transcription inverse, est médié par une ADN polymérase ARN dépendante ou transcriptase reverse, codée par le gène *pol*. Les rétrovirus codent aussi au minimum pour deux gènes additionnels. Le gène *gag* code pour les protéines du squelette, matrice, nucléocapside et capsid. Le gène *env* code pour les glycoprotéines d'enveloppe.

20 La transcription rétrovirale est régulée par des régions promotrices ou "enhancers", situées dans des régions hautement répétées ou LTR (*Long Terminal Repeat*) et qui sont présentes aux deux extrémités du génome rétroviral.

 Lors de l'infection d'une cellule, la polymérase fait une copie ADN du génome ARN ; cette copie peut alors s'intégrer dans le génome humain. Les rétrovirus ne tuent pas les cellules qu'ils infectent, mais au contraire améliorent souvent leur rapidité de croissance. Les rétrovirus peuvent infecter des cellules germinales ou

30

des embryons à un stade précoce ; ils peuvent dans ces conditions, intégrer la lignée germinale et être transmis par transmission mendélienne verticale. ce qui constitue la relation la plus étroite entre un hôte et son parasite. Ces virus endogènes peuvent dégénérer au cours des générations de l'organisme hôte et perdre leurs propriétés initiales. Cependant certains d'entre eux peuvent conserver tout ou partie de leurs propriétés ou des propriétés des motifs les composant, ou encore acquérir de nouvelles propriétés fonctionnelles présentant un avantage pour l'organisme hôte. ce qui expliquerait la préservation de leur séquence.

L'existence de motifs endogènes présentant de longs cadres de lecture ouverts et/ou soumis à une forte pression de sélection peut donc être indicatrice d'une fonction biologique préservée ou acquise, qui peut correspondre à un bénéfice pour l'organisme hôte. Ces séquences rétrovirales peuvent aussi subir, au cours des générations, des modifications discrètes qui vont être à même de réveiller certaines de leurs potentialités et engendrer ou favoriser des processus pathologiques. Il est apparu récemment nécessaire de faire le bilan et d'identifier ces séquences afin de pouvoir évaluer leur impact fonctionnel.

Les séquences rétrovirales endogènes humaines ou HERVs représentent une part importante du génome humain. Ces régions rétrovirales se présentent sous plusieurs formes :

- des structures rétrovirales endogènes complètes associant des motifs *gag*, *pol* et *env*, flanqués de séquences nucléiques répétées, qui montrent une analogie significative avec la structure LTR-*gag-pol-env*-LTR des rétrovirus infectieux,

- des séquences rétrovirales tronquées ; par exemple, les rétrotransposons sont privés de leur domaine *env* et les rétroposons ne possèdent pas les régions *env* et LTR.

Jusqu'à présent l'étude de ces régions du génome a été négligée chez l'Homme pour deux raisons essentielles :

- l'existence d'insertions/délétions qui peuvent décaler le cadre de lecture et de mutations qui modifient la séquence. Ces modifications entraînent des altérations de la structure et par conséquent de la fonction biologique de ces motifs.

- l'absence d'associations avérées avec des pathologies humaines.

La connaissance, récente de fragments significativement représentatifs du génome humain et une orientation des recherches vers une étude structure/fonction des motifs rétroviraux endogènes, ont permis de préciser l'intérêt de ces régions. L'implication de séquences endogènes tronquées ou complètes dans des pathologies chez l'animal est documentée ; par exemple leur association avec des processus tumoraux a été clairement mise en évidence (S.K. Chattopadhyay et coll., 1982, *Nature*, **295**, 25-31). Une recherche visant à préciser l'association ou l'influence des HERVs dans des pathologies humaines se justifie donc aujourd'hui.

10 Une classification des éléments HERV a été proposée (Tönjes R.R. et al., *AIDS & Hum. Retrovirol.*, 1996, **13**, S261-S267; A.M. Krieg et al., *FASEB J.*, 1992, **6**, 2537-2544). Elle est basée sur une homologie de ces séquences avec des rétrovirus isolés chez les animaux, à l'aide de sondes rétrovirales hétérologues. En effet, en général, les HERVs présentent relativement peu d'homologie avec des rétrovirus infectieux humains connus.

Les familles de classe I présentent une homologie de séquence avec les rétrovirus de mammifères de type C ; on peut citer notamment la superfamille ERI, proche du virus MuLV (*murine leukemia virus*) et du virus BaEV (*baboon endogenous virus*).

20 Les familles de classe II présentent une homologie de séquence avec les rétrovirus de mammifères de type B tel que le MMTV (*mouse mammary tumour virus*) ou les rétrovirus de type D tel que le SRV (*squirrel monkey retrovirus*).

D'autres familles ont également été décrites ; parmi celles-ci, on peut citer des HERVs qui présentent, de manière exceptionnelle, une homologie partielle avec HTLV-1 (RTVL-H) ou des virus de primates ; HRES-1, par exemple, présente une homologie de séquence avec des HTLVs.

25 Les programmes de très grand séquençage du génome humain permettent aujourd'hui de disposer d'un nombre significatif de nouvelles séquences rétrovirales. L'usage de logiciels de traitement de données permet d'identifier et d'analyser ces gènes. Dans ce contexte une recherche systématique portant sur l'ensemble des informations disponibles à ce jour a été engagée afin d'identifier de nouvelles séquences

30

rétrovirales endogènes humaines en fonction de certains critères d'analyse :

- présence de longs cadres de lecture ouverts conservés au cours de l'évolution de l'organisme hôte et pouvant laisser envisager une fonction biologique.

- analogie avec des séquences déjà caractérisées en dehors ou dans le
5 domaine des rétrovirus,

- localisation dans des régions de susceptibilité pour certaines pathologies ou à proximité de gènes essentiels, par exemple dans les domaines du cancer, de la régulation du système immunitaire ou dans certaines neuropathologies.

Les recherches effectuées par les Inventeurs, dans des bases de données de séquences leur ont permis d'identifier un ensemble de séquences ou de motifs
10 rétroviraux endogènes dont l'expression normale ou pathologique peut favoriser ou perturber un effet protecteur vis-à-vis de processus pathologiques, ou intervenir dans le déclenchement ou l'aggravation de pathologies.

La présente invention a pour objet un fragment d'acide nucléique
15 purifié, caractérisé en ce qu'il comprend tout ou partie d'une séquence codant pour une séquence rétrovirale endogène humaine, qui présente au moins des motifs rétroviraux de type *env*, répondant à la séquence SEQ ID NO:1 ou à une séquence présentant un niveau d'homologie avec ladite séquence SEQ ID NO:1 supérieur ou égal à 80% sur
plus de 190 nucléotides ou supérieur ou égal à 70 % sur plus de 600 nucléotides pour
20 les domaines de type *env*.

On entend par séquence homologue, aussi bien une séquence qui présente une identité complète ou partielle avec la séquence SEQ ID NO:1 précitée qu'une séquence qui présente une similarité partielle avec ladite séquence SEQ ID NO:1.

25 Selon un mode de réalisation avantageux dudit fragment, il présente à la fois des motifs rétroviraux correspondant à un domaine *env* et répondant à la séquence SEQ ID NO:1 et des motifs rétroviraux correspondant à un domaine *gag* et répondant à la séquence SEQ ID NO:2 ou à une séquence présentant un niveau d'homologie supérieur ou égal à 80 % sur plus de 190 nucléotides ou supérieur ou
30 égal à 70 % sur plus de 600 nucléotides pour les domaines de type *env* et un niveau d'homologie supérieur ou égal à 90 % sur plus de 700 nucléotides ou supérieur ou

égal à 70 % sur plus de 1200 nucléotides pour les domaines de type *gag*, lesquels motifs ne présentent aucune insertion ou délétion de plus de 200 nucléotides.

Lesdits fragments constituent une nouvelle famille de séquences rétrovirales endogènes humaines (famille HERV-7q) qui présente une homologie de
 5 séquence avec les rétrovirus MSRV, tels que décrits dans la Demande Internationale WO 97/06260 ; lesdits fragments selon la présente invention présentent :

- deux motifs nucléotidiques répétés de 711 pb (figure 3), présentant des signaux caractéristiques identifiés dans des LTRs (*Long Terminal Repeats*) : promoteurs de transcription de type boîtes TATAA ou CCAAT. Ces domaines répétés
 10 encadrent trois motifs déduits de type-*gag*, *pol* et *env* (figure 2).

- un motif de type *env* (positions 6965 nt - 9550 nt sur la séquence SEQ ID NO :3 ou sur la figure 1) qui contient un long cadre de lecture ouvert de 1620 nucléotides (positions 7874-9493 de la séquence ID NO:3 et figure 1), codant pour une protéine de séquence inédite de 540 acides aminés appelée envérine (figure 4 et
 15 SEQ ID NO:26) et fragment souligné de la figure 18. On retrouve à l'intérieur du domaine trans-membranaire de ce domaine *env*, un motif peptidique de type CKS-25/CKS-17 (figure 5), reconnu pour présenter des fonctions immunosuppressives sur les cellules lymphocytaires hôtes (M. Mitani et coll., 1987, *Proc. Natl. Acad. Sci. USA*, **84**, 237-240). Un domaine de type doigt de zinc (*zinc-finger*) **HX₃₋₄HX₂₂₋₃₃CX₂C** (Kulkolski et coll., 1992, *Mol. Cell. Biol.*, **12**, 2331-2338), que l'on retrouve
 20 dans des domaines de type intégrase est identifié dans un autre cadre de lecture. Ce domaine *env* particulier signe la caractéristique de nouveaux motifs rétroviraux endogènes.

- le motif (positions 3065 nt - 4390 nt sur la séquence SEQ ID NO:3)
 25 de type-*gag* codant pour des motifs protéiques selon la figure 6 (SEQ ID NO:58) (positions 3118-4198 de la SEQ ID NO:3) a été identifié grâce à des analogies avec des domaines *gag* connus. On retrouve, par exemple, la région d'homologie majeure **QX₃EX₇R** (Benit et coll., 1997, *J. Virol.*, **71**, 5652-5657). Le motif de fixation des acides nucléiques **CX₂CX₃₋₄HX₄C**, situé en position C-terminale, est identifié dans un
 30 autre cadre de lecture (Covey et coll., 1986, *Nucleic Acids Res.*, **14**, 623-633). En amont du domaine *gag* on détecte un motif de 182 nucléotides répété deux fois (figure

1).

- le domaine *pol* présente les consensus classiques d'une région *pol* de rétrovirus au niveau des domaines protéase, transcriptase reverse et RNase H. On retrouve dans *pol* un motif proche du consensus **LLDTGA** (Weber et coll., 1988, Science, **243**, 928-931). Les motifs **D** et **AF**, **LPQ** et **SP**, et **YVDD** (Xiong et Eickbush, 1990, EMBO J., **9**, 3353-3362), sont respectivement retrouvés dans les 3°, 4° et 5° boîtes d'homologie. Les motifs **YTDGSS** et **TDS** sont présents dans la région de la RNase H,

- les régions *gag* et *pol* pourraient être considérées comme jointives avec un passage de la région *gag* à la région *pol* par un décalage du cadre de lecture.

La présente invention englobe les séquences appartenant à la famille HERV-7q telle que définie ci-dessus (présence de la séquence SEQ ID NO:1 ou d'une séquence homologue ou présence à la fois des séquences SEQ ID NO:1 et SEQ ID NO:2) et notamment les séquences SEQ ID NO:3-22, 28 et 61 ; elle englobe également les séquences nucléiques complémentaires et les séquences inverses complémentaires des séquences précédentes ainsi que les fragments issus des régions codantes des séquences précédentes correspondant à un cadre glissant supérieur ou égal à 14 nucléotides ou leurs séquences complémentaires. (SEQ ID NO :37-57, 59-60 et 121-122).

Ces différents fragments peuvent avantageusement être utilisés comme amorces ou comme sondes (réactifs A) ; ils s'hybrident spécifiquement dans des conditions de forte stringence à une séquence de la famille HERV-7q.

Parmi ces fragments, on peut citer, de préférence les fragments suivants:

- un fragment de 182 nucléotides répété deux fois, situé en amont du domaine *gag* aux positions 2502-2611/2613-2865 de la SEQ ID NO:3 ;

Amorces et sondes spécifiques de la région *gag*

- une amorce G1F, sens, localisée dans la région amont du domaine *gag* de HERV-7q : 5' GGACCATAGAGGACACTCCAGGACTA 3' (SEQ ID NO:37);

- une amorce G1R, anti-sens, localisée dans la région 3' terminale du

domaine *gag* : 5' CCTCAGTCCTGCTGCTGGATCATCT 3' (SEQ ID NO:38)

- le fragment de 1505 nt amplifié par le couple G1F-G1R est utilisé afin de générer les sondes aptes à hybrider les différents produits d'amplification des PCR ;

- 5 - une amorce G2F, sens nichée : (SEQ ID NO:39)
 5' CCTCCAAGCAGTGGGAGGAAGAGAATT 3'
 - une amorce G2R, anti-sens nichée : (SEQ ID NO:40)
 5' CCTTCCCTGTGTTATTGTGGACATCATT 3'
 - une amorce G4F, sens nichée : (SEQ ID NO:41)
 10 5' GGAAGAAGTCTATGAATTATTCAATGATGT 3'
 - une amorce G3F, sens nichée: (SEQ ID NO:42)
 5' GGGACACAGAATCAGAACATGGAGATT 3'
 - une amorce G4R, anti-sens nichée : (SEQ ID NO:43)
 5' GCCTTCAGAAGAGTCAGGTGACAGAGA 3'
 15 - une amorce G5R, anti-sens nichée : (SEQ ID NO:44)
 5'GAGCCTCCAAAGTCCACTTGCCTGA 3'

Amorces et sondes spécifiques de la région *env*

- une amorce E1F, sens : (SEQ ID NO:45)
 5' GATTTCAGTATCTACTAGTCTGGGTAGAT 3'
 20 - une amorce E1R, anti-sens : (SEQ ID NO:46)
 5' CTAGGAAATCCAGCTAGTCCTGTCTCA 3'
 - le fragment de 2529 nt amplifié par le couple d'amorces E1F-E1R,
 est utilisé afin de générer les sondes aptes à hybrider les différents produits
 d'amplification des PCR.
 25 - une amorce E2F, sens : (SEQ ID NO:47)
 5' CCAAGACAGCCAACTTAGTTGCAGACAT 3'
 - une amorce E2R, antisens : (SEQ ID NO:48)
 5' GGACGCTGCATTCTCCATAGAACTCTT 3'
 - une amorce E3F, sens : (SEQ ID NO:49)
 30 5' GCAATACTACATACACAACCAACTCCCAA 3'

- une amorce E3R, anti-sens : (SEQ ID NO:50)
5' GGGGGAGGCATATCCAACAGTTAGTA 3'
- une amorce E4F, sens : (SEQ ID NO:51)
5' CCATCTACACTGAACAAGATTTATACACTT 3'
- 5 - une amorce E4R, anti-sens : (SEQ ID NO:52)
5' AATGCCAGTACCTAGTGCACCTAGCACT 3'
- une amorce E5F, sens : (SEQ ID NO:53)
5' CGAATACAACGTAGAGCAGAGGAGCTTCGAA 3'
- une amorce E6F, sens : (SEQ ID NO:54)
10 5' AGCCCAAGATGCAGTCCAAGACTAAGAT 3'
- une amorce E5R : (SEQ ID NO:55)
5' GCGTAGTAGAGTTGTGCAGCTGAGAT 3'
- une amorce ExF : (SEQ ID NO:56)
CCCTTACCAAGAGTTTCTATGGAGAAT
- 15 - une amorce ExR : (SEQ ID NO:57)
ACCGCTCTAACTGCTTCCTGCTGAATT

Tous les oligonucléotides sont conçus pour pouvoir générer une amorce sens et une amorce anti-sens par un décalage de la séquence de l'amorce de référence de 1 à 7 nucléotides vers le côté 5' ou vers le côté 3': la modification de la

20 séquence peut entraîner une modification de la taille de l'amorce de 1 à 7 nucléotides selon les cas. Les amorces choisies peuvent être optimisées selon les cas par un raccourcissement ou un allongement portant sur 1 à 9 nucléotides.

De manière préférée, l'hybridation, le clonage, le sous-clonage, l'obtention, la préparation et l'analyse des acides nucléiques, des peptides et des anti-

25 corps, le séquençage des acides nucléiques et des peptides, l'hybridation *in situ* et l'immunohistochimie sont réalisés dans les conditions décrites dans les ouvrages suivants :

- Current Protocols in Molecular Biology. Eds. F.M Ausubel, R. Brent & R.E Kingston et coll. Green Publishing associates and Wiley Interscience.
- 30 - Molecular Cloning: a laboratory manual. Eds. J. Sambrook, E.F. Fritsch & T. Maniatis. Cold Spring Harbor Laboratory Press. Cold Spring Harbor.

- The Practical Approach series. Eds. D. Rickwood & B.D. Ames. IRL Press and Oxford University Press. En particulier, antibodies I & II; DNA cloning I, II, III; Nucleic acid and protein sequence analysis; Nucleic acid hybridization; Nucleic acid sequencing ; Oligonucleotide synthesis; Protein purification applications;
- 5 Protein purification methods; Protein sequencing; Transcription and translation; Gels electrophoresis of nucleic acids; Gels electrophoresis of proteins; Genome analysis; HPLC of macromolecules; Human genetic diseases; Microcomputing in biology; Molecular neurobiology; Mutagenicity testing; Essential molecular biology I & II.
- Proteome research: New frontiers in functional genomics. Eds
- 10 M.R. Wilkins & coll.. Springer.

La séquence rétrovirale endogène humaine (SEQ ID NO:3). située sur le bras long du chromosome 7 correspond à la séquence HERV-7q ; elle présente 10,5 kb (fig. 1 et 2) et répond aux critères précédemment définis.

- La recherche de domaines présentant des similitudes, tout ou partie,
- 15 avec les régions *gag* et *env* de HERV-7q a abouti à l'identification de nouvelles séquences rétrovirales endogènes. Ces séquences peuvent présenter la structure d'un rétrovirus endogène complet comme la séquence rétrovirale endogène située à proximité du gène des sous-unités alpha et delta du récepteur des cellules-T, et dénommée en conséquence HERV-TcR ; à titre d'exemple la figure 7 montre la comparaison des
- 20 alignements nucléiques des domaines *gag* respectifs de HERV-7q et HERV-TcR (séquence HG12, SEQ ID NO:19). On trouve aussi des structures rétrovirales partielles. Ces domaines rétroviraux similaires à HERV-7q sont identifiées dans des séquences nucléiques indépendantes comme le montre leur localisation chromosomique. Des motifs nucléiques (appelés ici, HEx ou HGx et respectivement analogues à des domaines de type *env* ou *gag*) ressemblant aux domaines *env* ou *gag* de
- 25 HERV-7q ont été retrouvés, à l'aide des banques de données précitées :

- HE2 : chromosome 17 (SEQ ID NO:4),
- HE3 et HG3: chromosome 6 (SEQ ID NO:5 et 6),
- HE4 : chromosome X (SEQ ID NO:7),
- 30 - HE5 : chromosome X q22 (SEQ ID NO:8),
- HE6 et HG6 : chromosome 1 q23.3-q24.3 (SEQ ID NO:9 et 10),

- HE7 : chromosome 7 p15 (SEQ ID NO:11),
- HE8 et HG8 : chromosome 19 (SEQ ID NO:12 et 13),
- HE9 : chromosome X (SEQ ID NO:14),
- HE10 : chromosome X q13.1-21.1 (SEQ ID NO:15),
- 5 - HE11 et HG11 : chromosome 7 q21-22 (SEQ ID NO:16 et 17),
- HE12 et HG12, dans HERV-TcR : chromosome 14 q11.2 (SEQ ID NO:18 et 19),
- HE13 (SEQ ID NO:61) : chromosome 6 q24.1-24.3

La présente invention englobe également les fragments codants et
 10 non codants pour tout ou partie de l'envérine comprenant au moins 14 nucléotides et notamment les fragments codant pour la partie C-terminale de l'envérine, soit à partir de l'acide aminé 291, soit à partir de l'acide aminé 321, à compter de la première méthionine.

Ces fragments comprennent en particulier une zone critique où deux
 15 insertions de 12 nucléotides ont été caractérisées :

- une première insertion a été identifiée (séquence A), chez des individus de 2 groupes (malades et témoins). Cette insertion située entre les acides aminé 487 et 488, permet d'insérer le térapeptide VLQM. Une analyse comparative montre que cette insertion est identifiée dans une région homologue située dans la séquence
 20 HE13, appartenant à la famille HERV-7q. L'amplification de la séquence de type HE13, pourrait indiquer qu'il existe une altération de la séquence de l'envérine de HERV-7q, ce qui favoriserait l'amplification de la séquence contenue dans HE13. Cette observation permet aussi d'utiliser cette insertion comme élément spécifique d'amplification de séquences de type HE13.

25 Une deuxième insertion (séquence B) a été identifiée chez un patient présentant une SEP. L'insertion de 12 nucléotides est située au niveau de l'acide aminé 495 et code pour le térapeptide MQSM. Il est remarquable de constater que cette insertion est aussi identifiée dans une région homologue située dans HE13.

Séquence A: TAAACTACAAATGGTTCTTCAAATGGAGCCCA
 30 (SEQ ID NO:59)

Séquence B: GATGCAGTCCAAGATTGCAGTCCATGACTAAGA
(SEQ ID NO:60).

Ces observations mettent en évidence des modifications de la
séquence de l'envérine de type HERV-7q qui constituent la base d'une stratégie de
5 détection par amplification spécifique d'allèles (AS-PCR), permettant de détecter ces
différences dans une population et qui pourraient correspondre, soit à une muta-
tion/délétion associée à une certaine susceptibilité, soit à un polymorphisme, soit à une
mutation/délétion associée à une pathologie comme la sclérose en plaques.

Les alignements des domaines *env* (fig. 8) et *gag* (fig. 9) explicitent
10 les niveaux d'homologie observés entre les séquences décrites ci-dessus et les séquen-
ces homologues dans HERV-7q. Les analogies peuvent s'étendre aux motifs rétro-
viraux flanquants.

Une analyse des séquences étiquettes disponibles dans les banques
de données montre que des transcrits appartenant à certains des membres de cette
15 famille, en particulier HERV-7q, s'expriment essentiellement dans des tissus d'origine
fœtale ou placentaire.

Des séquences polypeptidiques générées par ces transcrits peuvent
donc être potentiellement produites et des fonctions ou activités biologiques peuvent
être envisagées, par analogie avec des polypeptides biologiquement actifs d'origine
20 virale ou rétrovirale ; par exemple, les motifs peptidiques de type CKS-17 (Haraguchi
et al., PNAS, 1995, 92, 5568-5571) (fig. 5) ou CKS-25 (Huang S.S et Huang J.S, J.
Biol. Chem. 1998, 273, 4815-4818), qui présentent des fonctions immunomodulatrices
sur les cellules lymphocytaires hôtes. Les différences de séquence observées et
d'éventuelles modifications normales ou pathologiques, sont en particulier, à l'origine
25 d'une modulation de la fonction.

HERV-7q représente le paradigme de la nouvelle famille de
séquences rétrovirales endogènes humaines ou de motifs rétroviraux endogènes.

HERV-7q et certaines des séquences rétrovirales endogènes apparte-
nant à sa famille, présentent un domaine de type *pol* analogue à des séquences rétro-
30 virales de type *pol* comme par exemple la région *pol* identifiée dans le rétrovirus
MSRV associé à la sclérose en plaques et décrit par H. Perron et al. (1997, *Proc. Natl.*

Acad. Sci. USA, **94**, 7583-7588 ; Demande Internationale PCT WO 97/06260).

Toutefois, les séquences selon la présente invention se distinguent des séquences rétrovirales exogènes infectieuses analogues à MSRV antérieurement décrites en ce que les séquences *gag* et *env*, selon l'invention sont significativement
5 différentes selon les critères précédemment définis et en fonction de certaines caractéristiques spécifiques, par exemple le long cadre de lecture ouvert du domaine *env* de HERV-7q ; elles seraient à même de permettre de signer une pathologie lorsqu'elles présentent des insertions, des délétions, des décalages de cadre de lecture ou des mutations.

10 En effet, les différences observées entre les séquences humaines de type HERV-7q, qui sont isolées d'individus réputés normaux et les séquences issues de certains échantillons d'origine pathologique, ne sont pas distribuées au hasard. Des comparaisons menées entre la région *gag* provenant de particules rétrovirales infectieuses (N° d'accèsion EMBL: A60168, A60200, A60201, A60171...) et la séquence
15 *gag* correspondante de HERV-7q (fig. 9), permettent d'observer que les mutations affectent préférentiellement des codons non-sens. Par exemple, deux codons non-sens dans HERV-7q sont remplacés par un codon arginine dans A60200, ce qui permet d'obtenir une séquence déduite de 109 acides aminés pour HERV-7q et de 166 acides aminés pour A60200. Les changements de base permettent en conséquence de prolon-
20 ger le cadre de lecture et de coder potentiellement pour des structures polypeptidiques de plus grande taille (figure 10).

De même, une séquence de type *env* provenant de particules rétrovirales infectieuses, présente une analogie significative avec le domaine *env* de HERV-7q (figure 11). Ces analogies marquées entre séquences rétrovirales exogènes et endogènes pourraient être à l'origine du déclenchement ou de l'aggravation de
25 certains processus pathologiques, en particulier de certaines maladies auto-immunes, comme la sclérose en plaques. A cet égard, on peut remarquer que certaines des séquences rétrovirales endogènes décrites dans l'invention se situent à proximité ou dans des régions réputées présenter une susceptibilité pour la sclérose en plaques : par
30 exemple HERV-7q et la région 7q21-22 du chromosome 7, de même pour HE12 et HG12 dans HERV-TcR et la région du gène codant pour les chaînes alpha et delta du

récepteur des cellules-T, HE2 et le chromosome 17, ou HE3, HE13 et HG3 et le chromosome 6, par exemple, les séquences HE11 et HG11, autour de la région 7q 21-22 ou encore HE4, HE5, HE6, HE9, HE10 ou HG10 sur le chromosome X. Ces séquences seraient donc à même de fournir des moyens de localisation ou
5 d'identification des gènes de prédisposition.

On n'observe aucune homologie significative avec des séquences rétrovirales endogènes déjà décrites, par contre, on peut relever une homologie limitée, permettant d'identifier une structure générale de domaine *env* ; toutefois, ladite homologie est inférieure aux critères définis selon l'invention entre les
10 domaines *env* de la séquence HERV-7q (SEQ ID NO :1) et de la séquence HERV-9 (figure 12). La figure 11 montre des homologies étendues entre la séquence HERV-7q avec une séquence rétrovirale exogène (N° d'accension EMBL : A60170).

Les séquences rétrovirales endogènes humaines appartenant à la famille de HERV-7q, peuvent protéger contre des agressions liées à l'environnement ou constituer un bénéfice pour l'individu. Cet effet bénéfique pourrait être une des
15 raisons possibles de la pression de sélection exercée sur certaines de ces séquences et du caractère potentiellement fonctionnel des structures protéiques déduites identifiées : par exemple le long cadre de lecture ouvert apte à coder pour une nouvelle protéine et correspondant au domaine *env* de HERV-7q.

Les séquences rétrovirales endogènes humaines appartenant à la famille de HERV-7q pourraient être associées par exemple, à des pathologies en relation avec les processus liés au cancer, aux neuropathologies à composante auto-immune ou à tout autre processus pathologique en association ou non avec des virus ou rétrovirus endogènes ou exogènes. Leur action pourrait porter sur la déclaration, l'aggravation, la
20 modification du calendrier d'apparition ou encore la protection vis à vis de la maladie.

Dans le contexte d'application à des pathologies autoimmunes (comme par exemple le lupus, le syndrome de Sjögren, la polyarthrite rhumatoïde, la sclérose en plaques...), on peut relever des analogies significatives entre les motifs rétroviraux endogènes identifiés et des motifs retrouvés dans des structures rétrovirales caractérisées chez des patients présentant des pathologies autoimmunes comme la
30 sclérose en plaques : par exemple des fragments de domaine *gag* (récemment dispo-

nibles dans les banques de données) provenant de particules rétrovirales infectieuses ou encore la séquence complète du domaine *pol* correspondant au virus MSRV associé à la sclérose en plaques. Ces motifs rétroviraux possèdent des analogies significatives avec les séquences endogènes homologues de type HERV-7q, ce qui permet
5 d'envisager une association directe ou indirecte avec des processus pathologiques, dont la sclérose en plaques, en association ou non avec MSRV.

L'intérêt de ces séquences dépasse le cadre des maladies auto-immunes. En dehors de l'importance générale des motifs rétroviraux dans le déclenchement ou l'aggravation d'un processus tumoral, bien montré en particulier dans les
10 modèles murins (H. Fan dans *The retroviridae*, 1994, ed. J.A. Levy, Plenum. New York, p. 313-353), ces séquences pourraient se retrouver à proximité ou au sein de gènes importants et en altérer l'expression : par exemple HERV-TcR et les gènes des sous-unités alpha et delta du récepteur des cellules T impliquées dans des perturbations de la fonction immunitaire.

15 La présente invention englobe, en outre, l'utilisation de séquences associées aux séquences de la famille HERV-7q pour la détection et/ou le pronostic de différentes maladies auto-immunes (neuropathologies, en particulier) ; ces séquences codent pour tout ou partie d'un facteur, dont la fonction, la régulation/dérégulation ou l'altération (polyadénylation, épissage alternatif) est associée à l'expression normale
20 ou pathologique ou à la régulation/dérégulation des motifs appartenant à la famille HERV-7q et correspondent à des transcrits ou des ADNc des séquences nucléotidiques codant pour des gènes situés dans des régions flanquantes ou encadrant des séquences rétrovirales de la famille HERV-7q.

On entend par région flanquante, toute région située à proximité
25 (incluse dans ou incluant) une séquence rétrovirale endogène appartenant à la famille HERV-7q, telle que définie ci-dessus, jusque et y compris les gènes immédiatement contigus et/ou situés à une distance ne pouvant excéder 120 kb.

Les Inventeurs ont maintenant trouvé que la présence des séquences rétrovirales telles que définies ci-dessus, perturbent l'expression ou altèrent la struc-
30 ture des séquences flanquantes définies ci-après.

Les transcrits desdites séquences flanquantes (et leurs fragments,

notamment ceux soulignés ou en italique dans les figures 14-16, 22-26, sont définies ci-après :

- à 1021 pb en amont de HERV-7q, on identifie une séquence rétrovirale endogène appelée RH7 (SEQ ID NO:62 et figure 22) ; cette séquence est située en 5' de la séquence HERV-7q ; dans la figure 22, la partie en italique correspond au début de la séquence HERV-7q ; la séquence RH7 est soulignée ; deux sites de polyadénylation putatifs sont en gras. Cette séquence SEQ ID NO:62 présente une homologie significative, sur plus de 6 kb, avec des séquences rétrovirales endogènes de type RGH (figure 13). Des séquences appartenant à cette famille s'expriment en particulier chez des patients présentant une arthrose rhumatoïde (Nakagawa et coll., (1997), *Arthritis. Rheum.*, 40, 627- 638). La présente invention inclut également des fragments de la séquence SEQ ID NO:62, comprenant entre 14 et 50 nucléotides (utilisation comme amorces), de préférence entre 14 et 25 nucléotides ou au moins 25 nucléotides (utilisation comme sonde), lesquels fragments présentent les caractéristiques suivantes : les 4 nucléotides de l'extrémité 3' sont différents des motifs correspondant de la séquence RGH2 (séquence du bas dans la figure 13, n° accession à GenBank : D110 18).

- à moins de 9 kb en amont de HERV-7q, on identifie la séquence RAM75 (SEQ ID NO:63 et figure 14) contenant les 24 exons codants (qui couvrent près de 41 kb), du gène de l'ATPase péroxysomale PEX1. PEX1, en association avec PEX6 est responsable de l'importation des protéines péroxysomales et de la stabilisation du récepteur PEX5. Une perturbation/altération affectant PEX1 est responsable de diverses neuropathologies comme le syndrome de Zellweger, l'adrénoleucodystrophie néonatale et la forme infantile de la maladie de Refsum (Reuber et coll., (1997), *Nature Genet.*, 17, 445- 448). On peut rappeler que la fonction principale des péroxysomes est associée au métabolisme des acides gras, en particulier par des processus de β -oxydation. Une altération du gène identifié dans la séquence RAM75 ou de son expression, par modification de la fonction des régions 5' et 3' régulatrices ou encore par modification des épissages ou des processus de polyadénylation, en particulier sous l'influence de motifs rétroviraux voisins, seraient à même de perturber l'expression ou la structure de l'ATPase et par conséquent de perturber une des

fonctions péroxy-somales, en particulier le métabolisme des lipides, en particulier myéliniques, avec des conséquences pour certaines pathologies, dont des neuropathologies, comme la sclérose en plaques ; les parties soulignées (figure 14) correspondent aux 24 exons codants.

5 La présente invention inclut également les fragments de la séquence SEQ ID NO:63, inclus dans les 24 exons codants précités et comprenant au moins 14 nucléotides.

L'analyse du profil d'expression (transcrits et protéines) de la séquence RAM75 (SEQ ID NO:63) est un bon indicateur du diagnostic différentiel des
10 neuropathologies à composante auto-immune.

A la figure 14, les exons codants sont soulignés. Les codons d'initiation et non-sens ainsi que les sites putatifs de polyadénylation sont en gras et soulignés.

- à 0.7 kb en aval de la séquence HERV-7q et sur près de 17 kb
15 (SEQ ID NO:64 et figure 15), on identifie la séquence nucléotidique RAV73, où l'on détecte des séquences étiquettes et des exons potentiels aptes à produire une ou plusieurs séquences polypeptidiques ; l'invention inclut également des fragments de cette séquence SEQ ID NO:64 compris dans les séquences étiquettes et les exons potentiels tels qu'ils apparaissent (parties soulignées) à la figure 15, lesquels
20 fragments comportent au moins 14 nucléotides.

- à 120 kb en amont de la séquence HG3, et sur 15 kb, se trouve la séquence nucléotidique RBP3 (SEQ ID NO:65 et figure 23), qui couvre l'extrémité 3'du gène codant pour un facteur de transcription de la famille Blimp-1 (SEQ ID NO:119 et figure 25), une protéine de 789 acides aminés qui est un répresseur de
25 l'expression du gène de l'interféron-béta (Keller et Maniatis, Genes Dev., (1991), 5, 868-879), qui est déjà associé à certaines pathologies malignes (Mock et coll., Genomics, (1996), 37, 24-28), et qui pourrait jouer un rôle dans la différenciation et la pathogenèse des cellules B. L'intérêt de l'association possible de la séquence rétrovirale endogène contenant les motifs HG3 et HE3 et de Blimp-1 est multiple, dans le
30 cas de pathologies, et en particulier la sclérose en plaques. Blimp-1 agit en particulier sur les cellules B dont on connaît la contribution dans les processus inflammatoires

associés à la sclérose en plaques. Blimp-1 est capable de bloquer l'induction virale du promoteur $INF\beta$ dont on connaît l'aptitude à réduire la fréquence des poussées et la progression lésionnelle chez des patients atteints de SEP. Une perturbation de l'expression ou de la structure de Blimp-1, en relation avec un élément rétroviral de type HERV-7q, est associée en conséquence à des neuropathologies ou à des maladies à caractère auto-immun, comme la sclérose en plaques ; cette séquence nucléotidique RBP3 (SEQ ID NO:65) contient des motifs nucléotidiques identifiés dans la séquence nucléique codant pour le gène Blimp-1 ; l'invention inclut aussi la détection des séquences ARNm de la protéine Blimp-1 (SEQ ID NO:119).

10 - la séquence rétrovirale endogène de type HERV-7q, contenant HE3 et HG3, se trouve située dans la région HI3 correspondant à un intron s'étendant sur plus de 46 kb (SEQ ID NO:66), d'un gène qui pourrait coder pour l'analogue d'APS (figure 24), une protéine de 275 acides aminés spécifique d'apoptose, surexprimée dans différents cellules en culture après déclenchement d'un processus apoptotique
15 (Hammond et coll., FEBS Lett., (1998), 425, 391- 395). L'intron se situe au niveau de l'acide aminé 231 d'APS. L'extrémité de HE3 est à plus de 12 kb de l'extrémité 5' de l'intron, alors que HG3 est situé à plus de 28 kb de l'extrémité 3' de l'intron. Des processus apoptotiques sont associés à la sclérose en plaques. En particulier, il a été décrit un processus apoptotique affectant des astrocytes et des oligodendrocytes en
20 présence d'une fraction purifiée de liquide céphalo-rachidien de patients atteints de sclérose en plaques (Ménard et coll., J. Neurol. Sci., (1998) , 154, 209- 221).

Enfin, il faut souligner que la région nucléique contenant HE3, HG3, HI3 et RBP3 est localisée au niveau du bras court du chromosome 6, en 6p21, qui est une région proposée de susceptibilité à la sclérose en plaques (The Multiple Sclerosis
25 Genetic Group, Nature Genet., (1996), 13, 469- 472).

L'interaction entre les séquences de type HERV-7q et les séquences flanquantes et l'importance de l'établissement d'un profil d'expression incluant une ou plusieurs des séquences précitées pour établir un diagnostic différentiel d'une neuro-pathologie apparaît encore plus du fait que l'on observe que les séquences HG12 et
30 HE12 sont situées dans une région intronique du gène codant pour les sous-unités alpha et delta des récepteurs des cellules T. Les récepteurs des cellules T sont impli-

qués dans les processus de régulation immunitaire et leur influence a été proposée dans le cas de maladies auto-immunes, dont la sclérose en plaques.

L'invention a également pour objet les transcrits générés à partir des séquences précitées ainsi que celles présentant éventuellement des modifications avec
5 les séquences de référence décrites dans l'invention lorsqu'ils sont exprimés chez certains patients.

En effet, les systèmes de régulation de l'expression des protéines rétrovirales de HERV-7q, qui sont présents dans les motifs de type LTR, pourraient influencer l'expression de gènes situés dans le voisinage chromosomique proche ou
10 éloigné et induire des perturbations à caractère immunologique et/ou neurologique. Par exemple la séquence rétrovirale endogène HERV-TcR, se trouve à proximité immédiate des gènes des sous-unités alpha et delta du récepteur des cellules-T précédemment décrit. Les motifs de type LTR pourraient aussi coder pour des superanti-
gènes (Acha-Orbea et Palmer, 1991, *Immunol. Today*, **12**, 356-361). D'une manière
15 générale des protéines rétrovirales de type HERV-7q ou apparenté, ou leurs formes tronquées ou partielles pourraient être impliquées dans des phénomènes de cytotoxicité ou de superantégenité, comme par exemple celles issues du long cadre de lecture ouvert identifié dans le domaine *env* (figure 4).

Des séquences du type des LTR 5' et 3' de HERV-7q, fortement
20 conservées sont concernées par de tels effets régulateurs. A titre d'exemple on décrit LTX, une séquence comparable à celle d'une LTR de HERV-7q (SEQ ID NO:67 et figure 16), et qui se trouve au cœur d'un intron de plus de 49 kb, mais à 2 kb du site 5' donneur, du gène de FMR2 associé au X-fragile et codant pour une protéine de 1311 acides aminés (figure 26). Les LTR modulent l'épissage alternatif (Kapitonov et
25 Jurka, (1999), *J. Mol. Evol.*, **48**, 248- 251), l'expression du gène, la fixation sur des protéines nucléaires (Akopov et coll., (1998), *FEBS Lett.*, **421**, 229- 233), ou permettent l'obtention d'un signal de polyadénylation alternatif (Goodchild et coll., (1992), *Gene*, **121**, 287- 294).

D'une manière générale, on peut remarquer l'existence de plusieurs
30 séquences rétrovirales endogènes de type HERV-7q (HE4, HE5, HE9, HE10), situées au niveau du chromosome X qui représente le chromosome associé au plus grand

nombre de pathologies.

A cet égard, on peut relever que des motifs rétroviraux issus de régions défectives sont aptes à présenter des fonctions biologiques: par exemple, la protéine d'enveloppe p15E issue de motifs rétroviraux défectifs, possède une activité
5 anti-inflammatoire et immunosuppressive (Snyderman et Ciancolo, 1984, *Immunol. Today*, **5**, 240-244).

Ces structures sont vraisemblablement à même de provoquer des brèches ou d'amplifier des dérégulations dans les processus de défense immunitaire. Certains des motifs des domaines *gag*, *env* et de type LTR peuvent être associés à une
10 fonction particulière ou peuvent contribuer à la fonction normale ou pathologique des domaines flanquants tels que définis ci-dessus (SEQ ID NO:62-67). Des recombinaisons avec un élément d'origine exogène, rétroviral ou non, peut donner lieu à la production de motifs nucléiques ou protéiques qui pourraient soit protéger, soit déclencher, ou favoriser ou aggraver une pathologie. De même, une structure rétro-
15 virale contenant des éléments rétroviraux endogènes selon l'invention seraient à même de provoquer un processus pathologique après passage par un cycle transitoire exogène puis réintégration dans une région sensible ou critique du génome humain.

On peut ainsi obtenir des profils d'expression (transcrits et éventuellement protéines) qui correspondent aux neuropathologies précitées.

20 De même, la combinaison de motifs appartenant à la famille de HERV-7q, ou d'éléments induits par des motifs appartenant à la famille de HERV-7q, avec des motifs d'origine ou induits de manière exogène seraient à même de pouvoir déclencher, ou aggraver un processus pathologique ou au contraire de favoriser une protection ou une rémission partielle ou une guérison totale et définitive.

25 La détection rendue possible des domaines de type HERV-7q, suggère des applications possibles à la fois au niveau prophylactique, du pronostic et du diagnostic : par exemple des approches immunologiques ou d'amplification génique permettant de comparer des individus normaux servant de référence avec des patients, seraient à même de favoriser le dépistage, d'améliorer la détection précoce de
30 la déclaration de la maladie et/ou de suivre l'évolution d'une pathologie chez des patients pouvant présenter une susceptibilité ou ayant déclaré la maladie ou encore

chez des individus considérés comme normaux, selon les critères cliniques actuels.

Les sondes nucléiques et immunologiques spécifiques, telles que définies, dans la présente invention sont à même de favoriser l'identification et la détection de motifs anormalement exprimés dans le cadre de pathologies associées au cancer, ou de neuropathologies en particulier auto-immunes, au premier rang
5 desquelles la sclérose en plaques.

La présente invention a également pour objet des séquences nucléiques hybrides, caractérisées en ce qu'elles comprennent des séquences ou motifs appartenant à la famille de HERV-7q, ou d'éléments induits par des motifs appartenant à la famille de HERV-7q, avec des motifs d'origine ou induits de manière
10 exogène (séquences rétrovirales exogènes) ; de telles séquences hybrides sont vraisemblablement à même de pouvoir déclencher, ou aggraver un processus pathologique ou au contraire de favoriser une protection ou une rémission partielle ou une guérison totale et définitive.

La présente invention a également pour objet un réactif de diagnostic pour la détection différentielle de séquences nucléiques endogènes humaines complètes ou partielles, présentant des motifs rétroviraux, sélectionnés parmi les séquences SEQ ID NO :1 et/ou SEQ ID NO :2, caractérisé en ce qu'il est sélectionné dans le groupe constitué par les séquences SEQ ID NO:1-22, 28, 37-57, 59-61 et 121-
20 122, les séquences nucléiques complémentaires et les séquences inverses complémentaires des séquences précédentes, par les fragments nucléotidiques capables de définir ou d'identifier les séquences SEQ ID NO:1 et/ou SEQ ID NO:2 et toute séquence flanquante ou les chevauchant ainsi que par les fragments issus des régions codantes des séquences SEQ ID NO:1-22 et 61, correspondant à un cadre glissant
25 supérieur ou égal à 14 nucléotides ou leurs séquences complémentaires, éventuellement marquées avec un marqueur approprié ainsi que par les séquences telles que définies aux figures 18-21.

Les séquences des sondes nucléiques, ribonucléiques et oligonucléotidiques utilisées seront choisies dans les régions *env* et *gag* ou leur régions
30 flanquantes : par exemple les oligonucléotides amorces pour HERV-7q, seront choisis dans les régions situées entre les nucléotides 3065 et 4390, les nucléotides 6965 et

9550 ou les nucléotides 2502-2865 de la SEQ ID NO:3, ainsi que dans toute séquence adjacente (amont ou aval) capable de permettre une amplification spécifique (figure 1).

Parmi les marqueurs appropriés, on peut citer, les isotopes radio-actifs, les enzymes, les fluorochromes, des marqueurs chimiques (biotine), les haptènes (digoxygénine) et les anticorps ou analogues de bases appropriées.

De manière préférée :

- ledit réactif est sélectionné parmi les séquences SEQ ID NO:37-57 et est apte à être utilisé comme amorce.

10 - ledit réactif est sélectionné parmi les séquences suivantes :

un fragment de 1505 nt amplifié par le couple d'amorces SEQ ID NO:37 et SEQ ID NO:38 (amorces G1F et G1R),

un fragment de 2529 nt amplifié par le couple d'amorces SEQ ID NO:45 et SEQ ID NO:46 (amorces E1F et E1R),

15 un fragment de 182 nucléotides répété deux fois, situé en amont du domaine *gag* aux positions 2502-2611/2613-2865,

des fragments codants ou non-codants pour tout ou partie de l'envérine, comprenant au moins 14 nucléotides et notamment les fragments codant pour la partie C-terminale de l'envérine, soit à partir de l'acide aminé 291, soit à partir de l'acide aminé 321, à compter de la première méthionine,

20 et est apte à être utilisé comme sonde.

La présente invention a également pour objet un procédé de détection rapide et différentiel des séquences nucléiques rétrovirales endogènes de type *env* ou *env* et *gag*, de leurs variants normaux ou pathologiques, par hybridation et/ou amplification génique, réalisé à partir d'un échantillon biologique, lequel procédé est caractérisé en ce qu'il comprend :

(a) une étape dans laquelle l'on met en contact un échantillon biologique à analyser avec au moins une sonde telle que définie ci-dessus et

(b) une étape dans laquelle on détecte par tout moyen approprié, le
30 ou les produits résultant de l'interaction séquence nucléotidique-sonde.

Conformément audit procédé, il peut comprendre :

* préalablement à l'étape (a) :

. une étape de préparation du tissu ou du liquide biologique concerné,

. une étape d'extraction de l'acide nucléique à détecter, et

5 . au moins un cycle d'amplification génique et

* postérieurement à l'étape (b) :

. une étape de comparaison des séquences nucléiques obtenues dans ledit échantillon biologique avec les séquences rétrovirales endogènes humaines selon l'invention par tout moyen approprié et notamment par séquençage, Southern-blot, 10 coupure de restriction, SSCP ou toute autre méthode permettant d'identifier une insertion ou une délétion ou encore une simple mutation entre les différentes séquences comparées.

Conformément à l'invention, les séquences rétrovirales endogènes humaines selon l'invention sont ainsi comparées aux séquences nucléiques présentes 15 dans l'échantillon biologique à analyser et permettent la détection de séquences homologues de patients atteints de pathologies, susceptibles de mettre en jeu une modification de leur génome.

De manière avantageuse, lesdites comparaisons géniques sont menées à partir d'ADN génomique provenant d'individus témoins et de patients.

20 Une amplification génique classique par PCR sera menée à l'aide d'amorces 5' -sens et 3' -antisens encadrant ou comprenant la zone à étudier (zone *env* ou zone *gag*).

Également de manière avantageuse, les séquences des sondes nucléiques, ribonucléiques et oligonucléotidiques utilisées sont choisies dans les 25 régions *env* et *gag* ou leurs régions flanquantes : par exemple les oligonucléotides amorces pour HERV-7q, seront choisis dans les régions situées entre les nucléotides 3065 et 4390 et les nucléotides 6965 et 9550, ainsi que dans toute séquence adjacente (amont ou aval) capable de permettre une amplification spécifique (figure 1), comme précisé ci-dessus. Elles sont de préférence sélectionnées dans le groupe constitué par 30 un fragment de 1505 nt amplifié par le couple d'amorces SEQ ID NO:37 et SEQ ID NO:38 (amorces G1F et G1R),

un fragment de 2529 nt amplifié par le couple d'amorces SEQ ID NO:45 et SEQ ID NO:46 (amorces E1F et E1R).

L'étape d'amplification génique est notamment réalisée à l'aide d'une des techniques d'amplification génique suivante : amplification par la Q β -
5 réplicase, PCR, LCR, ERA, CPR ou SDA.

La présente invention a également pour objet des séquences chimères, caractérisées en ce qu'elles sont constituées par un fragment de 17 à 40 nucléotides d'une séquence flanquante telle que définie ci-dessus associée à un motif rétroviral endogène de type HERV-7q comprenant entre 17 et 40 nucléotides, tel que
10 défini ci-dessus.

La présente invention a également pour objet un procédé de détection des transcrits, tels que définis ci-dessus, caractérisé en ce qu'il comprend :

- le prélèvement des ARN messagers provenant d'échantillons biologiques (tissus, cellules, fluides biologiques) témoins et d'un échantillon analogue
15 prélevé chez des patients et

- l'analyse qualitative et/ou quantitative desdits ARNm, par hybridation *in situ*, par dot-blot, Northern-blot, RNase mapping ou RT-PCR, à l'aide d'un réactif de diagnostic tel que défini ci-dessus.

La présente invention a également pour objet une méthode de détection
20 tion et/ou d'évaluation d'une sur-expression/sous-expression ou d'une modification d'au moins l'une des séquences ou fragments de séquences rétrovirales endogènes de type HERV-7q et/ou de leurs séquences flanquantes associées, caractérisée en ce qu'elle comprend :

- le dépôt sur un support approprié comme par exemple un filtre de
25 nylon, une lame de verre ou leur équivalent, de l'ADNc ou son équivalent provenant de clones, de produits de PCR obtenus à partir d'ADN génomique, de produits de RT-PCR provenant de transcrits ou encore de séquences oligonucléotidiques spécifiques, lesdites séquences d'ADN étant des séquences ou des fragments de séquences rétrovirales endogènes de type HERV-7q et/ou leurs séquences flanquantes, telles que
30 définies ci-dessus, constituées par les transcrits et ADNc des séquences génomiques, qui codent pour tout ou partie d'un facteur, dont la fonction, la régulation/dérégulation

ou l'altération est associée à l'expression normale ou pathologique ou à la régulation/dérégulation de motifs appartenant à ladite famille HERV-7q, ces séquences correspondant à des séquences nucléotidiques codant pour des gènes situés dans des régions flanquantes situées en amont et/ou en aval d'une séquence rétrovirale de ladite

5 famille HERV-7q et dont l'une des extrémités ne peut se trouver à une distance excédant 120 kb, et/ou une séquence chimère telle que définie ci-dessus.

- l'hybridation dudit support avec au moins une sonde marquée de manière adéquate obtenue, par exemple, par rétrotransposition d'un mélange d'ARN provenant de cellules, de tissus ou de liquides biologiques provenant de témoins

10 réputés normaux, de membres de populations ethniques différentes, de patients atteints de pathologies souvent associées à une expression de rétrovirus, comme les processus tumoraux, ou comme les maladies auto-immunes, et

- la détection des hybrides formés.

Selon un mode de mise en œuvre avantageux de ladite méthode,

15 ledit transcrit ou ADNc est sélectionné dans le groupe constitué par les séquences SEQ ID NO:62-67 et 119 et leurs fragments correspondant à un cadre glissant supérieur ou égal à 14 nucléotides ou leurs séquences complémentaires.

Selon un autre mode de mise en œuvre avantageux de ladite méthode, ledit support comprend en outre toute séquence rétrovirale endogène ou

20 exogène.

La méthode des puces à ADN (Bowtell, (1999), Nature Genet., 21, 25- 32), est utilisée pour évaluer la modification de l'expression de tout ou partie de certaines des séquences d'origine rétrovirale de type HERV-7q et des séquences flanquantes. Brièvement de l'ADN provenant de clones, de produits de PCR obtenus à

25 partir d'ADN génomique, de produits de RT-PCR provenant de transcrits ou encore de séquences oligonucléotidiques spécifiques, sont déposées sur un support, comme par exemple un filtre de nylon, une lame de verre ou leur équivalent. Les séquences nucléiques déposées couvrent les différents domaines rétroviraux décrits ci-dessus, ainsi que les séquences contiguës et les gènes flanquants. Afin de détecter d'éventuels

30 processus d'épissage alternatifs, des ADN spécifiques sont synthétisés par pas de 500- 600 nucléotides avec un chevauchement de 250- 300 nucléotides de part et d'autre.

Les épissages alternatifs déjà identifiés feront l'objet d'une synthèse spécifique. L'hybridation s'effectue à l'aide d'une sonde obtenue, par exemple, par rétrotransposition d'un mélange d'ARN provenant de cellules, de tissus ou de liquides biologiques provenant de témoins réputés normaux, de membres de populations ethniques différentes, de patients atteints de pathologies souvent associées à une expression de rétrovirus, comme les processus tumoraux, ou comme les maladies auto-immunes, dont la sclérose en plaques. Dans ce cas une fraction de μg et jusqu'à quelques μg d'ARNm ou jusqu'à quelques μg ou dizaines de μg d'ARN, selon la méthode utilisée et la taille de la puce d'ADN concernée, sont suffisants pour la synthèse de la sonde nucléique.

10 La sonde nucléique est marquée de manière adéquate, afin d'autoriser une détection ultérieure, comme par exemple par fluorescence ou par une méthode équivalente.

L'usage de sondes bi-, voire multicolores permet de préciser l'expression concertée de plusieurs gènes en parallèle, en bénéficiant de plus d'une normalisation précise. L'acquisition des résultats est effectuée automatiquement,

15 comme par exemple par un système de balayage laser ou son équivalent.

Deux types de puces à ADN sont conçues, d'une part des puces présentant un ensemble exhaustif de séquences, et d'autre part des puces à ADN spécifiques permettant un ciblage sur une application plus spécifique.

Par exemple, une séquence critique en ce qu'elle contiendrait une

20 différence portant sur une délétion, voire une mutation, est détectée à l'aide d'oligonucléotides spécifiques (Wang et coll., (1998), Science, 280, 1077- 1082). Le polymorphisme associé à une base ou à une mutation est détecté à l'aide de quatre oligonucléotides possédant une des quatre possibilités de séquence au niveau d'une base (A, C G ou T): pour chaque différence ponctuelle les 4 oligonucléotides sont

25 déposés et les intensités d'hybridation sont comparées. De plus, un épissage alternatif est détecté en utilisant des ADN correspondant à un seul exon effectif ou putatif : le gène est donc analysé exon par exon. Les puces à ADN concernent aussi, par extension, toute séquence rétrovirale endogène ou exogène, comme par exemple ERV-9, ERV-K, ERV-L, ERV-H, ERV-4, ERV-6, ERV-8, ERV-10, ERV-15, ERV-16, ERV-

30 17, ERV-18, ERV-21, ERV-24, ERV-33, ERV-34, ERV-36, ERV-40, ERV-42, ERV-MLN, ERV-FRD, ERV-FTD...), ainsi que toutes les séquences exoniques putatives

(identifiées par l'existence de séquences étiquettes et des transcrits correspondants) ou effectives, et qui sont situées de part et d'autre jusqu'à une distance de 120 kb des séquences rétrovirales endogènes de type HERV-7q.

L'étude comparative est menée entre un échantillon témoin et
5 l'échantillon à tester, dans une perspective prophylactique, diagnostique ou thérapeutique, comme par exemple: la détection précoce d'une modification de l'expression d'une des séquences, dans une cellule, un tissu ou un organisme, l'identification d'une séquence associée à une susceptibilité ou à une pathologie quelconque, le suivi de l'évolution de la pathologie, ou encore le suivi d'un traitement et l'évaluation de son
10 efficacité.

En dehors des applications déjà énoncées, l'intérêt de la méthode permet, d'une manière plus générale, de faire un bilan des variations constatées chez un individu, ce qui constitue en quelque sorte une carte d'identité, ce qui facilite une approche épidémiologique permettant d'établir de nouvelles corrélations entre un
15 profil particulier observé et une pathologie, en dehors de tout *a priori* concernant cette pathologie.

La présente invention a également pour objet un kit de détection et/ou d'évaluation d'une maladie auto-immune et notamment des neuropathologies à étiologie auto-immune, caractérisé en ce qu'il comprend outre les tampons nécessaires
20 à la mise en œuvre des procédés tels que définis ci-dessus :

- des réactifs A de diagnostic tels que définis ci-dessus, et
- des réactifs B constitués par les transcrits et ADNc des séquences génomiques, qui codent pour tout ou partie d'un facteur, dont la fonction, la régulation/dérégulation ou l'altération est associée à l'expression normale ou pathologique
25 ou à la régulation/dérégulation de motifs appartenant à ladite famille HERV-7q, ces séquences correspondant à des séquences nucléotidiques codant pour des gènes situés dans des régions flanquantes situées en amont et/ou en aval d'une séquence rétrovirale de ladite famille HERV-7q dont l'une des extrémités ne peut se trouver à une distance excédant 120 kb,

30 lesdits réactifs sont de préférence fixés sur un support approprié.
Selon un mode de réalisation avantageux dudit kit, lesdits réactifs B

sont sélectionnés dans le groupe constitué par les séquences SEQ ID NO:62-67 et 119 et leurs fragments correspondant à un cadre glissant supérieur ou égal à 14 nucléotides ou leurs séquences complémentaires, ainsi que les séquences représentées aux figures 13-17, 22-26.

5 La présente invention a également pour objet des produits de traduction, caractérisés en ce qu'ils sont codés par une séquence nucléotidique telle que définie ci-dessus.

La présente invention a également pour objet un peptide, caractérisé en ce qu'il est susceptible d'être exprimé à l'aide d'une séquence nucléotidique sélectionnée dans le groupe constitué par les séquences SEQ ID NO:1-22, 28 et 61, telles
10 que définies ci-dessus, selon les combinaisons offertes par l'usage des différents cadres de lecture possibles (voir également figures 18-21).

Ledit peptide englobe également les peptides ou polypeptides dérivés comprenant entre 5 et 540 aminoacides (SEQ ID NO:23-36 et SEQ ID NO:58 et
15 leurs fragments d'au moins 5 aminoacides) et notamment un fragment de 538 aminoacides, commençant à la première méthionine de la séquence SEQ ID NO:26 (envénérine).

Selon un mode de réalisation avantageux desdits peptides, ils sont notamment sélectionnés parmi les séquences SEQ ID NO:23-36, 58, notamment la
20 séquence SEQ ID NO:26 et ses fragments C-terminaux, soit à partir de l'acide aminé 291, soit à partir de l'acide aminé 321, à compter de la première méthionine.

Selon un autre mode de réalisation avantageux desdits peptides, ils sont obtenus à partir des séquences nucléiques telles que définies ci-dessus, dans lesquelles au moins un codon non-sens peut être remplacé par un codon codant pour
25 l'un des aminoacides suivants : Phe (F), Leu (L), Ser (S), Tyr (Y), Cys (C), Trp (W), Gln (Q), Arg (R), Lys (K), Glu (E) ou Gly (G).

L'invention englobe ainsi les peptides déduits ou les protéines déduites correspondant à tout ou partie des séquences nucléiques décrites dans l'invention, et présentant éventuellement des modifications avec les séquences de références décrites dans l'invention, lorsqu'ils sont exprimés chez certains patients. En
30 particulier, l'invention englobe les séquences complètes ou partielles obtenues selon

les 3 cadres de lecture sens et les 3 cadres de lecture inverses et complémentaires.
(voir figures 18-21)

De manière avantageuse, l'analyse de la structure du domaine env de
HERV-7q, appelé envérine, a permis de mettre successivement en évidence:

5 - un peptide signal N-terminal (région 1- 21) et deux domaines
transmembranaires (région 320-340; 455-477), responsables d'interactions avec des
motifs protéiques ou lipidiques membranaires,

 - un motif immunomodulateur de type CKS-17(Haraguchi et coll.,
(1995), 92, 5568- 5571)/ CKS-25. On peut noter à cet égard, la présence d'un motif
10 **RalD** à l'intérieur du peptide de type CKS-17/CKS-25 de HERV-7q et un motif **RvaD**
en position 363 qui correspondent au consensus W/RxxD, proposé pour le site actif
des TGF- β (Huang et al., J. Biol. Chem., 1997, 272, 27155- 27159), de puissants
facteurs associés à la croissance, à la différenciation et à la morphogenèse et qui sont
associés à de nombreuses pathologies humaines, comme les processus tumoraux
15 (Tang et coll., (1998), Nat. Med., 4, 802- 807) ou les maladies neurodégénératives
(Flanders et coll., (1998), Prog. Neurobiol., 54, 71- 85). Les peptides selon l'invention
contenant ces motifs peuvent avantageusement servir d'antagonistes en inhibant la
fixation des TGF- β sur leurs récepteurs naturels.

 - des motifs de N-glycosylation. La glycosylation des protéines
20 d'enveloppe des rétrovirus semble être directement associée à leurs propriétés fonc-
tionnelles, par exemple en influençant le nombre des déterminants disponibles dans les
cellules-T ou en favorisant la reconnaissance des antigènes par les cellules-T. La
glycosylation pourrait jouer un rôle dans la déclaration ou l'extension d'une patholo-
gie à incidence autoimmune. Les glycosylations sont nécessaires au maintien de la
25 conformation de certains épitopes, en particulier lors de la réalisation d'une protéine
d'enveloppe recombinante à fin de mise au point d'un réactif de diagnostic et pour
favoriser l'efficacité d'un éventuel vaccin. Positions 171, 210, 216, 236, 244, 283 et
411. Nombre prévu au hasard : 3.2

 - des sites de prénylation. La prénylation est un mécanisme essentiel
30 de la fixation à la membrane cellulaire et pour le ciblage de certaines protéines. Ce
processus de ciblage pourrait être essentiel pour l'élaboration d'agents thérapeutiques

spécifiques aptes à interférer dans la réalisation et la régulation du trafic de complexes cellulaires mettant en jeu des protéines impliquées dans les interactions, la croissance et les mouvements cellulaires. Positions 188 et 290. Nombre prévu au hasard : 1.8

- des sites de ciblage dans le réticulum endoplasmique. Ces sites
- 5 permettraient d'assurer le ciblage vers le réticulum endoplasmique afin d'effectuer les modifications nécessaires pour favoriser le franchissement membranaire. Positions 353 et 431. Nombre prévu au hasard : 0.2.

Par ailleurs, les Inventeurs ont montré qu'un certain nombre de peptides issus de la protéine env de HERV-7q (envérine) présentent une affinité/demi-
 10 vie élevées pour des allèles HLA de classe I. Une analyse par CADD a permis de sélectionner des peptides candidats, dont les meilleurs scores sont indiqués dans le Tableau I:

TABLEAU I

15	locali- sation	séquence	molécule HLA	score	Séquence n°
20	399	FLGEECCYYV	A-0201	7214	SEQ ID NO:68
	462	LLFGPCIFNL	A-0201	1792	SEQ ID NO:69
	189	CLPLNFRPYV	A-0201	1453	SEQ ID NO:70
	439	GLLSQWMPWI	A-0201	488	SEQ ID NO:71
	263	CLPSGIFFV	A-0201	5103	SEQ ID NO:72
	444	WMPWILPFL	A-0201	897	SEQ ID NO:73
25	252	IRWVTPPTQI	B-2705	3000	SEQ ID NO:74
	432	LRNTGPWGLL	B-2705	2000	SEQ ID NO:75
	158	LRTHTRLVSL	B-2705	2000	SEQ ID NO:76
	316	KRVPILPFVI	B-2705	1800	SEQ ID NO:77
30	25	CRCMTSSSPY	B-2705	1000	SEQ ID NO:78
	137	TRVHGTSSPY	B-2705	1000	SEQ ID NO:79
	124	AREKHVKEVI	B-2705	600	SEQ ID NO:80
	478	SRIEAVKLQM	B-2705	600	SEQ ID NO:81
	442	SQWMPWILPF	B-2705	500	SEQ ID NO:82
35	405	CYYVNQSGI	Kd	2400	SEQ ID NO:83
	346	FYYKLSQEL	Kd	2400	SEQ ID NO:84
	244	TYTTNSQCI	Kd	2400	SEQ ID NO:85
	291	SFLVPPMTI	Kd	1600	SEQ ID NO:86
	406	YYVNQSGIV	Kd	1200	SEQ ID NO:87
40	167	LFNTTLTGL	Kd	1152	SEQ ID NO:88
	463	LFGPCIFNL	Kd	960	SEQ ID NO:89
	253	RWVTPPTQI	Kd	480	SEQ ID NO:90
	449	LPFLGPLAAI	B-5102	2200	SEQ ID NO:91
	3	LPYHIFLFTV	B-5102	1210	SEQ ID NO:92

TABLEAU I (suite)

	locali- sation	séquence	molécule HLA	score	Séquence n°
5					
	331	GALGTGIGGI	B-5102	798	SEQ ID NO: 93
	321	LPFVIGAGVL	B-5102	550	SEQ ID NO: 94
	499	RRPLDRPAS	B-2705	600	SEQ ID NO: 95
10	194	FRPYVSIPV	B-2705	600	SEQ ID NO: 96
	383	RRALDLLTA	B-2705	600	SEQ ID NO: 97
	39	WRMQRPNGI	B-2705	600	SEQ ID NO: 98
	423	DRIQRRAEEL	B14	1800	SEQ ID NO: 99
	158	LRTHTRLVSL	B14	600	SEQ ID NO: 100
15	359	ERVADSLVTL	B14	540	SEQ ID NO: 101
	463	LFGPCIFNLL	Kd	1658	SEQ ID NO: 102
	345	QFYKLSQEL	Kd	1152	SEQ ID NO: 103
	443	QWMPWILPFL	Kd	691	SEQ ID NO: 104
	405	CYYVNQSGIV	Kd	500	SEQ ID NO: 105
20	474	NFVSSRIEAV	Kd	480	SEQ ID NO: 106
	221	GPLVSNLEI	B-5102	1320	SEQ ID NO: 107
	190	LPLNFRPYV	B-5102	726	SEQ ID NO: 108
	449	LPFLGPLAAI	B-5101	1144	SEQ ID NO: 109
	488	EPKMQSKTKI	B-5101	968	SEQ ID NO: 110
25	3	LPYHIFLFTV	B-5101	629	SEQ ID NO: 111
	125	REKHVKEVI	Kk	1000	SEQ ID NO: 112
	312	KPRNKRVPIL	B7	800	SEQ ID NO: 113
	378	VVLQNRRL	Db	792	SEQ ID NO: 114
	377	AVVLQNRRL	Db	660	SEQ ID NO: 115
30	321	LPFVIGAGV	B-5101	629	SEQ ID NO: 116
	304	DLYSYVISK	A3	540	SEQ ID NO: 117
	301	TEQDLYSYVI	Kk	500	SEQ ID NO: 118

Ce Tableau I indique une estimation de la demi-vie de dissociation d'un peptide de l'envérine avec un allèle du système HLA de classe I (les tables de coefficients de Parker: J. Immunol, (1994), 152, 163- 175). La localisation indique la position du premier acide aminé des peptides testés dans la séquence de l'envérine. Le code à une lettre est utilisé pour la séquence des acides aminés. Les scores autour de 500 ou supérieurs à 500 ont été retenus. A titre de comparaison, une analyse a été effectuée sur une concaténation de peptides (polypeptide de 4968 acides aminés) réputés pour fixer les molécules du complexe majeur d'histocompatibilité de classe I (Rammensee, Immunogenetics, (1995), 41, 178- 228): les dix meilleurs scores enregistrés pour des nonapeptides et le type HLA, A_0201 sont respectivement de 4984,

4047, 2406, 1267, 800, 705, 607, 591, 591 et 577.

Il ressort de ce Tableau I que certaines molécules du complexe majeur d'histocompatibilité de type I sont aptes à fixer des peptides issus de l'envérine, ainsi assimilés à des peptides d'origine virale ou tumorale, au niveau du réticulum endoplasmique. Les complexes formés au niveau du réticulum endoplasmique sont alors transportés à la surface cellulaire, ce qui entraîne la destruction de la cellule cible par les lymphocytes-T cytotoxiques. Les peptides identifiés comportent généralement 8 à 10 acides aminés. Des études ont montré que certains allèles du système HLA de classe I sont ainsi associées à certaines pathologies, en particulier à caractère auto-immun, comme HLA-B27 avec la spondylarthrite ankylosante ou HLA-B51 avec la maladie de Behçet.

Un peptide apte à fixer une molécule particulière de classe I est par conséquent apte à fonctionner comme un épitope de cellule-T.

En conséquence, la présente invention inclut également les fragments 399-471 et 244-271 de l'envérine qui regroupent avantageusement plusieurs épitopes de forte affinité pour différents haplotypes du système HLA de classe I. L'usage de tout ou partie de ces polypeptides est en conséquence apte à favoriser une augmentation du répertoire des cellules-T, en permettant une meilleure efficacité de la réponse immunitaire dans le cadre des différentes stratégies immunothérapeutiques, prophylactique ou vaccinales). Ces peptides pourront être avantageusement délivrés par exemple par l'usage, de vecteurs viraux, de particules virales ou synthétiques, de lipopeptides, d'adjuvants classiques, d'acides nucléiques nus ou adsorbés sur des particules, ou de liposomes.

Au sens de la présente invention, les peptides peuvent être chimiquement ou biochimiquement modifiés ; certaines des acides aminés peuvent être remplacés par un acide aminé analogue, selon les critères classiques d'homologies (A ou G ; S ou T ; I, L ou V ; F, Y ou W ; N ou Q ; D ou E).

La présente invention a également pour objet des compositions immunogènes ou vaccinales, pour la protection contre les maladies auto-immunes, notamment chez les sujets à risque, caractérisée en ce qu'elle comprend au moins un peptide comprenant au moins un motif de type CKS et/ou au moins un peptide

constitué par un motif présentant une affinité avec l'un des haplotypes du système HLA de classe I ou de classe II et un véhicule pharmaceutiquement acceptable.

Selon un mode de réalisation avantageux de ladite composition, ledit motif est sélectionné dans le groupe constitué par les peptides tels que définis dans le

5 Tableau I ci-dessus.

Selon un autre mode de réalisation avantageux de ladite composition, ledit peptide présente la séquence suivante :

séquence CKH: LQNRALDLLTAERGGTcI**FLGEECCYYV**
(SEQ ID NO:120).

10 Il est remarquable de constater au niveau de la position 380 de la protéine envérine, la contiguïté des motifs de type CKS-17 (souligné) et du peptide présentant le score le plus élevé (en gras ; voir peptide en position 399 dans le Tableau I, SEQ ID NO:68) dans la séquence CKH.

L'activation clonale des sous-groupes de lymphocytes, par exemple
15 de lymphocytes cytotoxiques, par les peptides du Tableau I et par extension leurs homologues, est bloquée par des manœuvres usuelles d'immunothérapie comme par exemple la sérothérapie et la vaccination.

L'association de deux séquences ou des séquences analogues au peptide CKH (SEQ ID NO:120), est à même d'entraîner un processus synergique dans
20 la réponse immunitaire, qui pourrait mettre en jeu des voies de signalisation et d'activation complémentaires, aptes à moduler l'activation lymphocytaire.

La vaccination concerne la production d'anticorps dirigés contre les peptides du tableau I, selon les règles de l'art et selon les méthodes de libération contrôlées par implants artificiels ou cellulaires mettant en œuvre une composition
25 telle que définie ci-dessus et par usage des manœuvres de thérapie génique, comme par exemple par expression des séquences nucléiques codant pour les peptides du Tableau I. En conséquence l'invention a également pour objet des compositions immunogènes ou vaccinale caractérisée en ce qu'elles comprennent un vecteur incluant au moins une séquence nucléique codant un peptide tel que défini dans le
30 Tableau I, éventuellement associée à une séquence codant un motif de type CKS-17.

La sérothérapie concerne l'utilisation d'anticorps neutralisants

produits à partir des peptides du Tableau I et leurs homologues.

Les produits protéiques générés par les séquences rétrovirales endogènes ou produits parallèlement peuvent avantageusement être caractérisés par des micro-méthodes d'analyse et de quantification des peptides et des protéines:

5 HPLC/FPLC ou équivalent, électrophorèse capillaire ou équivalent, techniques de microséquençages (méthode d'Edman ou équivalent, spectrométrie de masse...).

L'invention a également pour objet des anticorps dirigés contre l'un ou plusieurs des peptides décrits ci-dessus et leur utilisation soit pour la mise en œuvre d'une méthode de détection *in vitro*, notamment différentielle de la présence d'une

10 telle séquence chez un individu, soit pour la préparation d'une composition apte à être utilisée en sérothérapie dans les neuropathologies à composante auto-immune.

Lesdits anticorps sont avantageusement des anticorps polyclonaux ou monoclonaux obtenus par une réaction immunologique d'un organisme humain, mammifères, oiseaux ou autres espèces vis-à-vis des protéines, telles que définies ci-

15 dessus.

La présente invention a pour objet un procédé de dépistage immunologique différentiel de séquences rétrovirales endogènes humaines de la famille HERV-7q normales ou pathologiques, caractérisé en ce qu'il comprend la mise en contact d'un échantillon biologique avec un anticorps selon l'invention, la lecture

20 du résultat étant révélée par un moyen approprié, notamment EIA, ELISA, RIA, fluorescence.

A titre d'illustration, une telle méthode de diagnostic *in vitro* selon l'invention comprend la mise en contact d'un échantillon biologique prélevé chez un patient, avec des anticorps selon l'invention et la détection à l'aide de tout procédé

25 approprié, notamment à l'aide d'anti-immunoglobulines marquée, des complexes immunologiques formés entre les protéines produites normalement ou pathologiquement et les anticorps.

Des anticorps monoclonaux ou polyclonaux, produits à partir d'antigènes correspondants à des peptides de synthèse, de polypeptide ou protéines

30 recombinants, permettent de suivre l'expression des peptides ou protéines produits normalement ou pathologiquement. L'analyse est de préférence effectuée par ELISA,

ou équivalent, Western-blot ou équivalent, ou par immunohistochimie.

Les peptides ou protéines, issus des séquences rétrovirales endogènes ou dont l'expression est associée à l'expression de ces séquences rétrovirales endogènes, sont recherchés et identifiés.

5 La présente invention a également pour objet un procédé d'identification et de détection de motifs rétroviraux endogènes, anormalement exprimés dans le cadre de pathologies associées au cancer, ou de neuropathologies en particulier auto-immunes, au premier rang desquelles la sclérose en plaques, caractérisé en ce qu'il comprend l'analyse comparée des séquences extraites d'un échantillon biologique
10 avec les séquences selon l'invention.

La présente invention a également pour objet l'application des séquences nucléiques ou des séquences protéiques selon l'invention au diagnostic, au pronostic, à l'évaluation de la susceptibilité génétique, à toutes maladies humaines induites, innées ou acquises en particulier celles à composantes cancéreuses, auto-
15 immunes et/ou à incidence neurologique, comme la sclérose en plaques, les syndromes associés et les maladies neurodégénératives où intervient tout ou partie des séquences nucléiques selon l'invention et des formes endogènes ou exogènes apparentées.

La présente invention a également pour objet des séquences
20 nucléiques hybrides, caractérisées en ce qu'elles comprennent des séquences ou motifs nucléiques selon l'invention, combinés avec des séquences ou motifs d'origine endogène ou d'origine ou induits de manière exogène.

La présente invention a, en outre, pour objet un vecteur recombinant de clonage ou d'expression, caractérisé en ce qu'il comprend une séquence nucléique
25 conforme à l'invention.

Des manœuvres thérapeutiques peuvent être envisagées par usage de certaines des séquences nucléiques contenues dans HERV-7q et les séquences de la même famille ou des structures polypeptidiques déduites ou par utilisation de peptides ou protéines, ou d'anticorps spécifiques.

30 Conformément à l'invention, tout ou partie des séquences nucléiques rétrovirales endogènes de type HERV-7q, peut être utilisée à usage de vecteur ou

d'éléments de vecteurs à vocation thérapeutique, en particulier les séquences LTR et la région *gag* (SEQ ID NO :2, 21 et 22)..

L'intérêt de telles séquences réside, dans l'innocuité du vecteur ainsi formé, dans la possibilité d'une insertion spécifique ciblée dans une région bien définie par une stratégie analogue à la recombinaison homologue, dans le ciblage cellulaire, éventuellement transitoire dans le cas d'une expression placentaire chez la femme. Un autre aspect concerne la possibilité d'associer aux gènes d'intérêts les motifs rétroviraux biologiquement actifs (peptides immunomodulateurs, tels que représentés aux séquences SEQ ID NO 68-118, ci-après, peptide fusogène...).

La présente invention a également pour objet des animaux transgéniques, caractérisés en ce qu'ils comprennent tout ou partie d'une séquence de type HERV-7q (SEQ ID NO:1-22 et 61).

Le Tableau II ci-après établit les correspondances entre les numéros des séquences telles qu'elles apparaissent dans la liste de séquences et le nom des différentes séquences.

TABLEAU II

SEQ ID NO :	DÉSIGNATION
1	Acide nucléique : 7 env
2	Acide nucléique : gag
3	Acide nucléique : HERV-7q
4	Acide nucléique : HE2
5	Acide nucléique : HE3
6	Acide nucléique : HG3
7	Acide nucléique : HE4
8	Acide nucléique : HE5
9	Acide nucléique : HE6
10	Acide nucléique : HG6
11	Acide nucléique : HE7
12	Acide nucléique : HE8
13	Acide nucléique : HG8
14	Acide nucléique : HE9
15	Acide nucléique : HE10
16	Acide nucléique : HE11
17	Acide nucléique : HG11
18	Acide nucléique : HE12
19	Acide nucléique : HG12

SEQ ID NO:	DÉSIGNATION
20	Acide nucléique : R1
21	Acide nucléique : R1F
22	Acide nucléique + protéine env déduite : HERV-7q
23	Fragment de protéine env déduite selon SEQ ID NO :22
24	Fragment de protéine env déduite selon SEQ ID NO :22
25	Fragment de protéine env déduite selon SEQ ID NO :22
26	Protéine : envérine
27	Fragment de protéine env déduite selon SEQ ID NO :22
28	Acide nucléique + protéine déduite de gag : HERV-7q
29	Fragment de protéine gag déduite selon SEQ ID NO :28
30	Fragment de protéine gag déduite selon SEQ ID NO :28
31	Fragment de protéine gag déduite selon SEQ ID NO :28
32	Fragment de protéine gag déduite selon SEQ ID NO :28
33	Fragment de protéine gag déduite selon SEQ ID NO :28
34	Fragment de protéine gag déduite selon SEQ ID NO :28
35	Protéine env : cadre de lecture 1
36	Protéine gag
37	Acide nucléique : G1F (amorces)
38	Acide nucléique : G1R (amorces)
39	Acide nucléique : G2F (amorces)
40	Acide nucléique : G2R (amorces)
41	Acide nucléique : G4F (amorces)
42	Acide nucléique : G3F (amorces)
43	Acide nucléique : G4R (amorces)
44	Acide nucléique : G5R (amorces)
45	Acide nucléique : E1F (amorces)
46	Acide nucléique : E1R (amorces)
47	Acide nucléique : E2F (amorces)
48	Acide nucléique : E2R (amorces)
49	Acide nucléique : E3F (amorces)
50	Acide nucléique : E3R (amorces)
51	Acide nucléique : E4F (amorces)
52	Acide nucléique : E4R (amorces)
53	Acide nucléique : E5F (amorces)
54	Acide nucléique : E6F (amorces)
55	Acide nucléique : E5R (amorces)
56	Acide nucléique : ExF (amorces)
57	Acide nucléique : ExR (amorces)
58	Protéine gag
59	Acide nucléique : Séquence A (séquence d'insertion)
60	Acide nucléique : Séquence B (séquence d'insertion)
61	Acide nucléique : HE13
62	Acide nucléique : RH7

SEQ ID NO:	DÉSIGNATION
63	Acide nucléique : RAM75
64	Acide nucléique : RAV73
65	Acide nucléique : RBP3
66	Acide nucléique : HI3
67	Acide nucléique : LTX
68	Peptide Tableau I
69	Peptide Tableau I
70	Peptide Tableau I
71	Peptide Tableau I
72	Peptide Tableau I
73	Peptide Tableau I
74	Peptide Tableau I
75	Peptide Tableau I
76	Peptide Tableau I
77	Peptide Tableau I
78	Peptide Tableau I
79	Peptide Tableau I
80	Peptide Tableau I
81	Peptide Tableau I
82	Peptide Tableau I
83	Peptide Tableau I
84	Peptide Tableau I
85	Peptide Tableau I
86	Peptide Tableau I
87	Peptide Tableau I
88	Peptide Tableau I
89	Peptide Tableau I
90	Peptide Tableau I
91	Peptide Tableau I
92	Peptide Tableau I
93	Peptide Tableau I
94	Peptide Tableau I
95	Peptide Tableau I
96	Peptide Tableau I
97	Peptide Tableau I
98	Peptide Tableau I
99	Peptide Tableau I
100	Peptide Tableau I
101	Peptide Tableau I
102	Peptide Tableau I
103	Peptide Tableau I
104	Peptide Tableau I
105	Peptide Tableau I

106	Peptide Tableau I
SEQ ID NO:	DÉSIGNATION
107	Peptide Tableau I
108	Peptide Tableau I
109	Peptide Tableau I
110	Peptide Tableau I
111	Peptide Tableau I
112	Peptide Tableau I
113	Peptide Tableau I
114	Peptide Tableau I
115	Peptide Tableau I
116	Peptide Tableau I
117	Peptide Tableau I
118	Peptide Tableau I
119	Acide nucléique : BLIMP-1
120	Peptide : CKH
121	Acide nucléique : F645 (amorce)
122	Acide nucléique : PS5D (amorce)

Otre les dispositions qui précèdent, l'invention comprend encore d'autres dispositions, qui ressortiront de la description qui va suivre, qui se réfère à des exemples de mise en œuvre du procédé objet de la présente invention ainsi qu'aux
5 dessins annexés, dans lesquels :

- Figure 1. Séquence nucléique humaine HERV-7q, dont l'analyse et le traitement permettent de caractériser une nouvelle structure rétrovirale endogène. Les régions nucléiques répétées de type R1 et R2 et les domaines *gag*, *pol* et *env* sont soulignés. Les domaine de type *gag* et *env* sont en italiques. La région homologue à
10 une partie 3' non-codante de Rab7 est doublement soulignée.

- Figure 2. Cartographie de la région rétrovirale endogène humaine HERV-7q. La partie haute de la figure correspond à une région anonyme du génome humain située sur le bras long du chromosome 7. On peut identifier les domaines répétés (1), *gag* (2), *pol* (3) et *env* (4) de HERV-7q. La région *env* C-terminale (4.3) se
15 prolonge en amont en un long cadre de lecture ouvert (4.2). Le domaine 4.1, correspond à la région N-terminale du domaine *env*.

- Figure 3. Comparaison des séquences nucléiques répétées situées aux bornes de HERV-7q. Les régions nucléiques répétées 5'(haut) et 3'(bas), sont

comparées et les bases identiques sont indiquées par deux points.

- Figure 4. Séquence déduite présentant un cadre de lecture ouvert, dans le domaine de type-*env* de HERV-7q selon la règle du plus long cadre de lecture ouvert.

5 - Figure 5. Séquences autour du domaine CKS-17 identifiées dans différents domaines *env* déduits de la famille de HERV-7q et comparaison avec des motifs CKS-17 de référence.

1) HE2 - 2) HERV-7q - 3) N° d'accès à GenBank: M85205 - 4) HE7 - 5) HE9 - 6) CKS-17: le motif peptidique doué de propriétés immunomodulatriques est souligné - 7) gp20 de rétrovirus de type-D (SRV-Pc).

- Figure 6. Séquence déduite possible du domaine de type-*gag* identifié dans HERV-7q établie selon la règle du plus long cadre de lecture ouvert. X et / correspondent respectivement à un codon non-sens et à un décalage de cadre de lecture. La séquence soulignée correspond au début du domaine *pol*.

15 - Figure 7. Comparaison des régions nucléiques couvrant la région *gag* de HERV-7q (haut) et HERV-TcR (bas) et leurs régions flanquantes. Les bases identiques sont spécifiées par deux points.

- Figure 8. Exemple d'alignements nucléiques du domaine de type *env* de HERV-7q avec des domaines de type *env* similaires présents dans des séquences rétrovirales endogènes humaines de la même famille. Les codons non sens sont soulignés : 1) HERV-7q - 2) HE2 - 03) HE3 - 04) HE4.

- Figure 9. Alignements nucléiques entre le domaine *gag* de HERV-7q et les domaines correspondants appartenant à la même famille. Comparaison avec des fragments de domaines *gag* isolés d'agents rétroviraux infectieux. Séquences d'origine rétrovirale infectieuse: N° d'accession dans la banque de données EMBL : 1) A60168 - 2) A60201 - 3) A60200 - 4) A60171. Séquences rétrovirales endogènes humaines: 5) HERV-7q - 6) HG11 - 7) HG3. Les chiffres indiqués dans les séquences endogènes, correspondent au nombre de nucléotides insérés afin d'optimiser l'alignement avec les séquences de type *gag* identifiées dans des rétrovirus d'origine infectieuse.

- Figure 10. Alignement d'un motif *gag* protéique déduit (haut)

appartenant à un rétrovirus infectieux (N° d'accèsion EMBL : A60200) avec le motif *gag* protéique déduit (bas) identifié dans HERV-7q. Les codons non-sens sont en gras et soulignés. Les acides aminés identiques sont spécifiés par 2 tirets. Un tiret indique une délétion ou un acide aminé homologue.

- 5 - Figure 11. Alignement d'un motif *env* (haut) appartenant à un rétrovirus infectieux (N° d'accèsion EMBL : A60170) avec le motif *env* (bas) identifié dans HERV-7q. Les nucléotides homologues sont spécifiés par deux points et les délétions par un tiret.

- 10 - Figure 12. Comparaison entre le domaine *env* de HERV-7q (haut) et le domaine *env* de HERV-9 (bas). L'homologie de 66 % se limite à la région 3' du domaine *env* de HERV-7q et HERV-9, respectivement entre les nucléotides 8976 nt et 9500 nt de HERV-7q et les nucléotides 2898 nt et 3465 nt de HERV-9 (N° d'accèsion à GenBank : X57147). De nombreuses insertions/délétions sont aussi observées.

- 15 - Figure 13. Homologie entre une partie de la séquence du transcrit codant pour RH7 (haut, SEQ ID NO:62) et un motif de RGH2 (bas - N° d'accèsion à GenBank: D11018).

- 20 - Figure 14. Identification de la séquence du transcrit codant pour RAM75 (SEQ ID NO:63), correspondant au gène d'une ATPase de type PEX1. Les exons codants sont soulignés. Les codons d'initiation et non-sens ainsi que les sites putatifs de polyadénylation sont en gras et soulignés. La région en italique correspond au début de la séquence rétrovirale endogène RH7.

- 25 - Figure 15. Séquence du transcrit codant pour RAV73 (SEQ ID NO:64), située à 0.7 kb en aval de HERV-7q ; les séquences nucléiques aptes à coder pour un ou plusieurs polypeptides sont soulignées.

- Figure 16. Comparaison entre la séquence LTR 3' (haut) de HERV-7q et la séquence intronique LTX (SEQ ID NO:67), située dans le gène FMR2, associé au X-fragile (bas).

- 30 - Figure 17. Mise en évidence de modifications sur la séquence nucléotidique (ID NO:3), chez des patients atteints de SEP. Les bases modifiées, chez au moins un patient, sont soulignées. Les amorces utilisées sont en italiques

(séquences SEQ ID NO:121 et 122). L'ATG d'initiation et le codon non-sens sont en gras.

- Figure 18. Partie codante *env* de la séquence HERV-7q (séquence ID NO:3), avec 3 cadres de lecture.

5 - Figures 19, 20, 21. Présentation séparée de la protéine env selon les 3 cadres de lecture.

- Figure 22. Séquence nucléique contenant la séquence rétrovirale RH7 située en 5' de la séquence HERV-7q. La séquence en italique correspond au début de la séquence HERV-7q. La séquence RH7 est soulignée. Deux sites de poly-
10 adénylation putatifs sont gras.

- Figure 23. Séquence du transcrit codant pour RBP3 contenant des motifs nucléotidiques identifiés dans la séquence nucléique codant pour le gène Blimp-1.

- Figure 24. Séquence du transcrit codant pour APS.

15 - Figure 25. Séquence du transcrit codant pour Blimp-1 ; la partie codante est soulignée ; les codons d'initiation et de terminaison sont en gras.

- Figure 26. Séquence du transcrit codant pour FMR2. La partie codante est soulignée. Les codons d'initiation et non-sens sont en gras.

Il doit être bien entendu, toutefois, que ces exemples sont donnés
20 uniquement à titre d'illustration de l'objet de l'invention, dont ils ne constituent en aucune manière une limitation.

EXEMPLE 1 : Détection, par amplification génique, d'une séquence nucléique appartenant à un domaine de type *gag* ou *env* selon l'invention, dans un échantillon d'ADN génomique d'origine humaine ou de mammifères.

25 L'amplification génique s'effectue à partir d'ADN génomique isolé à partir du sang. Un traitement anticoagulant est effectué avec 1 ml d'une solution de citrate (pour un litre : 4,8 g de d'acide citrique, 13,2 g de citrate de sodium, 14,7 g de glucose) pour 6 ml de sang frais. Après centrifugation de 20 ml de sang pendant 15 mn à 13.0000 g, le surnageant est éliminé et la fraction enrichie en globules blancs est

transférée dans un nouveau tube, puis recentrifugée dans les mêmes conditions que précédemment. La fraction enrichie en globules blancs est resuspendue dans un tampon d'extraction (10 mM Tris-HCl, 0,1 M EDTA, 20 µg/ml de RNase pancréatique traitée afin d'éliminer les DNases, 0,5 % SDS, pH 8,0), puis incubée pendant 1 heure
5 à 37°C. La protéinase K est ajoutée à une concentration finale de 100 µg/ml. La suspension des cellules lysées est incubée à 50°C durant 3 heures sous agitation périodique, puis traitée par un volume égal de phénol équilibré par du Tris-HCl 0,5 M, pH 8,0. L'émulsion formée est placée sur une roue pendant une heure, puis centrifugée à 5000 g pendant 15 mn à température ambiante. La solution aqueuse est traitée déprotéinisée par une triple extraction phénolique afin d'obtenir un niveau de purification
10 correspondant à un rapport final d'absorbance A260/A280 supérieur à 1,75. La fraction aqueuse est précipitée par 0,2 vol. d'acétate de sodium 10 M et 2 vol. d'éthanol. L'ADN est alors soit prélevé avec l'extrémité d'une pipette pasteur recourbée, soit centrifugé à 5000 g pendant 5 mn à température ambiante. L'ADN ou le culot d'ADN
15 est lavé deux fois par de l'éthanol à 70 %, puis repris dans 1 ml de TE pH 8,0 afin d'être élué sous agitation douce pendant 12 à 24 heures.

Des oligonucléotides spécifiques des séquences endogènes décrites selon l'invention sont choisis pour amplifier la région *gag* ou *env* des régions rétrovirales endogènes décrites selon l'invention. L'ADN génomique étudié provient de
20 patients présentant des pathologies comme la sclérose en plaques et d'individus réputés sains.

Les ADN polymérases thermostables utilisées ont été choisies pour leur grande fidélité lors du processus d'amplification, comme la Vent_r ADN polymérase (Biolabs) ou équivalent, et sont utilisées selon les conditions préconisées par le
25 fournisseur.

La stratégie d'amplification utilise selon les cas une simple PCR, ou une PCR nichée ou semi-nichée.

Oligonucléotides utilisés pour amplifier la région *gag* :

- amorce G1F, sens, localisée dans la région amont du domaine *gag*
30 de *HERV-7q* (SEQ ID NO:37),
- amorce G1R, anti-sens, localisée dans la région 3' terminale du

domaine *gag* (SEQ ID NO:38),

Le fragment de 1505 nt amplifié par le couple G1F-G1R : 1505 nt est utilisé afin de générer les sondes aptes à hybrider les différents produits d'amplification des PCR.

- 5 - amorce G2F, sens nichée (SEQ ID NO:39),
- amorce G2R, anti-sens nichée (SEQ ID NO:40),
- amorce G4F, sens nichée (SEQ ID NO:41),
- amorce G3F, sens nichée (SEQ ID NO:42),
- amorce G4R, anti-sens nichée (SEQ ID NO:43),
- 10 - amorce G5R, anti-sens nichée (SEQ ID NO:44),

Oligonucléotides utilisés pour amplifier la région *env* de HERV-7q :

- amorce E1F, sens (SEQ ID NO:45),
- amorce E1R, anti-sens (SEQ ID NO:46),

15 Le fragment de 2529 nt amplifié par le couple d'amorces E1F-E1R, est utilisé afin de générer les sondes aptes à hybrider les différents produits d'amplification des PCR.

- amorce E2F, sens (SEQ ID NO:47),
- amorce E2R, antisens (SEQ ID NO:48),
- amorce E3F, sens (SEQ ID NO:49),
- 20 - amorce E3R, anti-sens (SEQ ID NO:50),
- amorce E4F, sens (SEQ ID NO:51),
- amorce E4R, anti-sens (SEQ ID NO:52),
- amorce E5F, sens (SEQ ID NO:53),
- amorce E6F, sens(SEQ ID NO:54)
- 25 - amorce E5R(SEQ ID NO:55).
- amorce ExF (SEQ ID NO:56)
- amorce ExR (SEQ ID NO:57)

La PCR est réalisée à partir de 50 à 200 ng d'ADN génomique. Les conditions de PCR sont celles préconisées par le fournisseur. Les conditions cycliques d'amplification sont réalisées dans 50 µl : une dénaturation de 94°C pendant 1 min., 30 une hybridation de 70°C pendant 1 min., et une élongation à 72 °C pendant 1 à 2 min.,

selon les fragments amplifiés. Après 35 cycles, une réaction terminale est menée à 72°C pendant 10 min. Le séquençage automatique des échantillons amplifiés est réalisé à l'aide d'un séquenceur Applied Biosystems de type ABI 377 ou autre modèle comparable, selon les protocoles fournis par le constructeur.

5 Dans le cas d'une PCR nichée ou semi-nichée, les mêmes conditions expérimentales sont utilisées, à la seule différence que l'échantillon d'ADN génomique est remplacé par 5 à 10 µl du produit d'amplification issu de la première PCR.

Deux amplifications indépendantes sont réalisées à partir du même échantillon. Une réaction de contrôle est réalisée en remplaçant l'échantillon d'ADN
10 par de l'eau afin de détecter d'éventuels contaminants.

EXEMPLE 2 : Détection par amplification génique d'une séquence nucléique selon l'invention dans un échantillon biologique d'ADN génomique prélevé chez des patients présentant une pathologie candidate déclarée ou la suspicion de cette pathologie.

15 Le protocole d'amplification est le même que dans l'exemple 1, mis à part l'origine de l'échantillon qui provient de patients présentant une pathologie candidate. Un échantillon d'ADN génomique réputé normal est systématiquement intégré dans l'ensemble des échantillons pathologiques amplifiés puis analysés.

Les produits de PCR sont séparés sur un gel d'agarose à 1,5 %, puis
20 transférés en présence de soude 0,4 N sur une membrane de nylon chargé. Une hybridation est réalisée avec une sonde spécifique correspondant aux fragments de PCR amplifiés soit par les couples G1F-G1R soit par le couple E1F-E1R. La sonde est marquée par incorporation de dUTP-digoxygénine selon le protocole du fournisseur (Boehringer Mannheim). L'hybridation est effectuée dans un tampon d'hybridation
25 (5XSSC, 50 % formamide, 0,1 % lauroyl-sarcosine, 0,02 % SDS, 2 % de réactif de blocage Boehringer) pendant une nuit à 42°C. Le Southern est lavé 2 fois 5 min. à température ambiante dans une solution de 2XSSC, 0,1% SDS. Puis un lavage à haute stringence est effectué à deux reprises pendant 15 min. à 55°C dans une solution
0,1XSSC, 0,1 % SDS. L'hybridation est révélée selon le protocole du fournisseur
30 (Boehringer Mannheim), en présence d'un substrat chimioluminescent de la phosphatase alcaline, de type CSPD ou CDP-STAR. Le filtre est révélé après une exposition

de 15min. à 60 min.

Une analyse par SSCP ("*single strand conformation polymorphism*") permet de détecter des modifications discrètes de la séquence des fragments amplifiés par PCR. La PCR est menée en présence de dCTP marqués au P³². L'échantillon à analyser est dénaturé à 95°C pendant 10 min., en présence de tampon de charge, puis immédiatement chargé sur un gel de polyacrylamide à 10%, contenant 7.5% de glycérol. La migration s'effectue à 4°C à 8-10 W. Le gel est séché puis autoradiographié.

Les fragments de PCR susceptibles de présenter une altération de leur séquence nucléotidique sont séquencés selon l'exemple 1.

Une hybridation à l'aide d'un oligonucléotide spécifique (17 mers à 20 mers) correspondant à la région nucléotidique modifiée permet d'identifier les échantillons présentant une modification identique (méthode ASO). Brièvement le southern est hybridé avec un oligonucléotide marqué distalement soit au P³², soit en présence de digoxygénine (selon le protocole de Boehringer Mannheim) puis lavé dans des conditions stringentes à 65°C dans une solution 6XSSC, 0.05% pyrophosphate de sodium.

Par exemple, un séquençage nucléotidique automatique a été réalisé sur six fragments de PCR, provenant de 5 patients atteints de SEP et un témoin réputé normal, et qui ont été amplifiés à partir des amorces F645: CTTCAAACAACAACCAGGAGG (SEQ ID NO:121) (située à 26 nucléotides en amont de la méthionine d'initiation de l'envérine) et PS5D: TTGGGGAGGTTGGCCGACGA (SEQ ID NO:122) (située à 6 nucléotides en aval du codon non-sens de l'envérine. Des modifications de la séquence de l'envérine ont été observés sur l'ADN de certains des patients (figure 17).

EXEMPLE 3 : Détection d'une protéine selon l'invention dans un échantillon biologique.

- Préparation d'une fraction protéique purifiée de liquide céphalo-rachidien de patients atteints de SEP

Après un traitement à 56°C pendant 30 min, et élimination des immunoglobulines sur une colonne de protéine G HiTrap (Pharmacia), l'équivalent de

10 ml de LCR est déposé sur une colonne de DEAE Sepharose CL-6B (Pharmacia). L'élution est réalisée en Tris-HCl 20 mM pH 8,8, et un gradient de 0 à 0,4 M de NaCl, puis la fraction est dialysée 2 fois contre du tampon phosphate-NaCl (PBS). Après concentration sur Ultrafree-MC (Millipore), la fraction est déposée sur une colonne de
5 Superose 12 (FPLC Pharmacia) et éluée en présence de PBS. Après séparation par électrophorèse en gel de polyacrylamide-SDS, et électro-transfert sur une membrane d'Immobilon-P (Millipore), les bandes protéiques sont soumises à une hydrolyse trypsique ménagée.

- Analyse de la fraction protéique par spectrométrie de masse

10 Les peptides digérés en présence de trypsine, sont analysés par la méthode de MALDI-TOF, qui permet l'analyse de peptides présents en mélange. (COTTRELL J.S., Pept. Res., 1997, 7, 115-124). Les peptides caractérisés en fonction de leur masse sont comparés aux protéines et aux protéines associées selon l'invention.

15 **EXEMPLE 4 : Détection d'anticorps spécifiques anti-domaine *env* de HERV-7q.**

L'identification d'un long cadre de lecture ouvert au sein de la séquence *env* de HERV-7q, a permis de déterminer une séquence protéique déduite SEQ ID NO:22 et 35 et figures 18-20 d'une région dudit gène.

20 Les séquences de protéines déduites des séquences ID NO:22, 35 et des figures 18-20 sont positionnées comme suit par rapport à la figure 1 ou à la séquence ID NO:3 :

SEQ ID NO:22 (cadre de lecture 1) et figure 19 : début de la séquence codante : position 7874, fin de la séquence codante 1^{er} codon non-sens (position 9493)

25 SEQ ID NO:35 : début de la séquence codante : position 7874, fin de la séquence codante 1^{er} codon non-sens (position 9493) (cadre de lecture 1)

Figure 19 : début de la séquence codante : position 6970, fin de la séquence codante 1^{er} codon non-sens (position 9493) (cadre de lecture 1)

30 Figure 20 : début de la séquence codante : position 6971, la fin du cadre de lecture est décalée selon le cas de 1, 2 ou 3 codons

Figure 21 : début de la séquence codante : position 6972, la fin du

cadre de lecture est décalée selon le cas de 1, 2 ou 3 codons

Différents peptides correspondant à tout ou partie des SEQ ID NO:22 (voir SEQ ID NO:23-27 et 35) ont été synthétisés par génie génétique afin de tester leur spécificité antigénique vis-à-vis de séra ou de tissus de patients atteints de SEP, par exemple. Brièvement, tout ou partie de la région env de HERV-7q est sous clonée dans les vecteurs pQE30, 31 et 32. Les vecteurs pQE30, 31 et 32 contiennent en 5' du multi-site de clonage les séquences consensuelles pour la transcription (le promoteur fort du bactériophage T5, 2 opérateurs de l'opéron lactose), la traduction (un site d'accrochage ribosomal synthétique). De même, pQE30, 31 et 32 possèdent en 3', le terminateur de transcription du phage λ ainsi qu'un codon "Stop" pour la traduction. L'expression de la protéine s'effectue après transformation dans *E. coli* M15. Le plasmide pQE30, 31 et 32 possèdent en amont du site de polyclonage la séquence codante pour une suite de 6 histidines présentant une affinité pour les ions nickel. Cet enchaînement permet la purification de la protéine chimérique exprimée, par adsorption sur une résine constituée d'un ligand chélatant, l'acide nitrilotriacétique (NTA), chargé de 4 ions nickel (résine NI-NTA, Qiagen).

La transformation s'effectue par électroporation ou traitement au chlorure de calcium. Par exemple, une colonie d'*E. coli* M15 est incubée dans 100 ml de milieu LB contenant 250 μ g de kanamycine, sous agitation à 37°C jusqu'à l'obtention d'une DO^{600} de 0,5. Après une centrifugation de 5 minutes à 2000g à 4°C, le culot bactérien est repris dans 30 ml de solution TFB1 (100 mM de chlorure de rubidium, 50 mM de chlorure de manganèse, 30 mM d'acétate de potassium, 10 mM $CaCl_2$, 15% glycérol, pH 5.8), à 4°C pendant 90 minutes. Après une centrifugation de 5 minutes à 2000g à 4°C, le culot bactérien est repris dans 4 ml de solution TFB2 (10 mM de chlorure de rubidium, 10 mM de MOPS, 75 mM $CaCl_2$, 15% de glycérol, pH 8). Les cellules peuvent être gardées à -70°C par aliquot de 500 μ l. 20 μ l de la ligation et 125 μ l de cellules compétentes sont mélangés et placés dans la glace 20 minutes. Après un choc thermique de 42°C pendant 90 secondes, les cellules sont agitées 90 minutes à 37°C dans 500 ml de milieu Psi-broth (milieu LB complémenté par 4 mM de $MgSO_4$, 10mM de chlorure de potassium). Les cellules transformées sont étalées sur des boîtes LB-agar complémentées par 25 μ g/ml de kanamycine, et 100 μ g/ml

d'ampicilline, et les boîtes sont incubées une nuit à 37°C.

Les clones potentiellement recombinants sont repiqués de manière ordonnée sur un filtre de nylon déposé sur une boîte LB-agar complémentée par 25 µg/ml de kanamycine et 100 µg/ml d'ampicilline. Après une nuit à 37°C, les clones
5 recombinants sont repérés par hybridation de l'ADN plasmidique avec la sonde nucléotidique amplifiée par PCR avec le couple d'amorces selon SEQ ID NO:45 et SEQ ID NO:46.

Une colonie indépendante, contenant l'insert, est inoculée à 20 ml de milieu LB complémentée par 25 µg/ml de kanamycine et 100 µg/ml d'ampicilline.
10 Après une nuit à 37°C sous agitation, 500 ml de même milieu sont incubés au 1/50° par cette préculture jusqu'à l'obtention d'une D0⁶⁰⁰ de 0,8, puis 1 à 2 mM final d'IPTG est ajouté. Après 5 heures, les cellules sont centrifugées 20 minutes à 4000 g.

Une partie du culot cellulaire est repris dans 5 ml de tampon de sonication (50 mM de phosphate de sodium pH 7,8, 300 mM NaCl) puis placé dans la
15 glace. Après une rapide sonication, les cellules sont centrifugées 20 minutes à 10000 g. Une partie du culot cellulaire est repris dans 10 ml d'une solution 30 mM Tris/HCl-20% sucrose pH8. Les cellules sont incubées 5 à 10 minutes sous agitation, après adjonction de 1 mM EDTA. Après une centrifugation de 10 minutes à 8000 g à 4°C, le culot est repris dans 10 ml de 5 mM de MgSO4 glacé. Après 10 minutes dans
20 la glace sous agitation, les cellules sont centrifugées 10 minutes à 8000 g à 4°C.

Le culot est repris par 5 ml/g dans du tampon A (6 M GuHCl (chlorhydrate de guanidine), 0,1M phosphate de sodium, 0,01M Tris/HCl, pH 8), 1 heure à température ambiante. Le lysat est centrifugé 15 minutes à 10000 g à 4°C, et le surnageant est complémenté par 8 ml de résine Ni-NTA, prééquilibrée dans du tampon
25 A. Après 45 minutes à température ambiante, la résine est coulée dans une colonne, lavée par 10 fois le volume de la colonne par du tampon A puis par 5 fois le volume de la colonne par du tampon B (8 M urée, 0,1 M phosphate de sodium, 0,01 M Tris/HCl, pH 8). La colonne est lavée par du tampon C (8 M urée, 0,1M phosphate de sodium, 0,01 M Tris/HCl, pH 6,3) jusqu'à ce que l'A280 soit inférieur à 0,01. La
30 protéine recombinante est éluée par 10 à 20 ml de tampon D (8 M urée, 0,1 M phosphate de sodium, 0,01 M Tris/HCl, pH 5,9) puis par 10 à 20 ml de tampon E (8 M

urée, 0,1 M phosphate de sodium, 0,01 M Tris/HCl, pH 4,5), puis par 20 ml de tampon F (6 M HCl, 0,2 M acide acétique). Après une analyse en SDS-PAGE, la ou les fractions purifiées contenant la protéine chimérique ont permis l'obtention d'anticorps chez le lapin. Les anticorps obtenus sont testés par Western-blot après révélation par
5 un anticorps secondaire couplé à la phosphatase alcaline.

Des anticorps sont obtenus de la même manière, à partir de peptides synthétisés chimiquement selon la technique de Merrifield (G. Barany and B. Merrifield, 1980, dans *The peptides*, 2, 1-284, E. Gross et J. Meienhofer, Academic Press, New York).

10 Les anticorps spécifiques obtenus sont utilisés à fin de détection de l'expression sérique ou tissulaire de tout ou partie des séquences rétrovirales endogènes selon l'invention, dans les cas normaux et pathologiques.

Les protéines d'origine sérique ou tissulaire, sont séparées sur gel d'acrylamide-SDS puis transférées sur un filtre de nitrocellulose à l'aide d'un appareil
15 Novablot 2117-2250 (LKB). Le transfert est effectué sur une feuille de Hybond C-extra (Amersham) en utilisant un tampon CAPS 100 mM pH 11, méthanol, eau (V/V/V: 1/1/8) contenant 1 mM de CaCl_2 . Après un transfert de 1 heure à 0,8 mA/cm², la feuille est saturée une heure à température ambiante dans du PBS-0,5 % gélatine. La feuille est mise en présence de l'anticorps spécifique à la concentration de
20 1/1000 dans du PBS-0,25 % gélatine. Au bout de 2 heures, le filtre est lavé 3 fois 15 minutes dans du PBS-0,1 % de Tween-20, puis le filtre est incubé 30 minutes en présence d'un anticorps secondaire couplé à la phosphatase alcaline (Promega), dilué au 1/7500 dans du PBS-0,25 % gélatine. Après trois lavages dans du PBS-0,1 % de Tween-20, le filtre est équilibré dans un tampon (100 mM de Tris-HCl pH 9,5, 100
25 mM de NaCl, 5 mM de MgCl_2). La révélation est effectuée en présence de 45 µl de NBT à 75 mg/ml et 35 µl de BCIP à 50 mg/ml, pour 10 ml de tampon de phosphatase alcaline.

Les protéines chimériques obtenues par génie génétique, sont utilisées aussi à fin de tests d'activité biologique, comme par exemple pour le test
30 d'activité biologique du peptide de type CKS-17 identifié dans le domaine *env* de

HERV-7q (figure 5).

EXEMPLE 5 : Obtention de sondes ribonucléiques codant pour les séquences *env* de HERV-7q.

Les fragments de PCR obtenus sont sous clonés dans le plasmide
5 PGEM 4Z (Promega) qui possède de par et d'autre de son site de polyclonage, les séquences promotrices pour les ARN polymérase SP6 et T7.

La méthode de compétence utilisée est l'électroporation. Le plasmide et le fragment de PCR sont hybridés dans un rapport de 50 ng de vecteur (coupé à Sma I) pour 100 ng de fragment de PCR (rendu à bout franc par traitement par le fragment
10 de Klenow de l'ADN polymérase). L'incubation a lieu une nuit à 22°C, dans le tampon de ligation (66 mM Tris-HCl pH 7,5, 5 mM MgCl₂, 1 mM dithioerythritol, 1 mM ATP) en présence de 1u. de T4 ADN ligase puis est arrêtée par dénaturation 10 minutes à 65°C. Parallèlement, la souche d'*E. Coli* JM 105 estensemencée une nuit à 37°C dans du milieu LB. Cette préculture est diluée au 1/500 et placée à 37°C jusqu'à
15 une DO⁶⁰⁰ égale à 1. Pour la suite du mode opératoire les cellules seront toujours conservées au froid. Après une centrifugation de 5 minutes à 3500 g à 4°C, le culot cellulaire est resuspendu dans 1/4 vol. d'eau glacée ultra-pure. Cette étape est répétée 5 à 6 fois. Puis le culot est resuspendu dans 1/4000 vol. d'eau; 10 % de glycérol stérile sont ajoutés permettant la conservation des cellules électrocompétentes, par aliquots
20 de 10 µl à 20°C. A 50 µl de cellules électrocompétentes est ajouté 1 µl de la ligation ; le tout est soumis à une décharge électrique de 12,5 kV/cm, appliquée pendant 5,8 ms. Les cellules sont rapidement remises en suspension dans le milieu SOC, incubées 1 heure à 37°C, puis étalées, en présence de 2% X-Gal dans du diméthylformamide, et 10 mM d'IPTG, sur une boîte de gélose LB-agar supplémentée en ampicilline (100
25 µg/ml). Après une nuit à 37°C, les clones blancs potentiellement recombinants, sont repiqués de manière ordonnée sur une boîte LB/ampicilline et parallèlement sur un filtre de nylon déposé sur une boîte LB/ampicilline. Ces deux boîtes sont incubées une nuit à 37°C. Les clones recombinants sont alors repérés par hybridation avec une sonde nucléique amplifiée par PCR avec le couple d'amorces selon SEQ ID NO:45 et
30 SEQ ID NO:46 et marquée à la digoxygénine.

Les clones recombinants sont cultivés dans 50 ml de milieu

LB/ampicilline (100 µg/ml) en agitation pendant une nuit à 37°C. Après une centrifugation à 3500 g pendant 15 minutes à 4°C, le culot bactérien est repris dans 4ml de tampon P1 (50 mM Tris-HCl, 10mM EDTA, 400 µg/ml RNase A, pH 8) et 4ml de tampon P2 (200 mM NaOH, 1% SDS). Le mélange est incubé à température ambiante pendant 5 minutes. Après adjonction de 4ml de tampon P3 (2.55 M d'acétate de potassium, pH 4,8) le mélange est centrifugé à 12000 g pendant 30 minutes à 4°C. Le surnageant est appliqué sur une colonne Qiagen-type 100, prééquilibrée avec 2 ml de tampon QBT (750 mM NaCl, 50 mM MOPS, 15% éthanol, pH 7). la colonne est lavée avec 2 fois 4ml de tampon QC (1M NaCl, 50 mM MOPS, 15 % éthanol, pH 7) et l'ADN est élué avec 2ml de tampon QF (1,2 M NaCl, 50mM MOPS, 15 % éthanol, pH 8). L'ADN est précipité avec 0,8 vol. d'isopropanol, et centrifugé à 12000 g à 4°C pendant 30 minutes. Le culot est lavé avec de l'éthanol à 70 % glacé. puis l'ADN plasmidique est repris par 2 fois 150 µl de tampon TE.

Les sondes ribonucléiques sont utilisées comme sondes spécifiques, en particulier pour la détection des transcrits exprimés par les séquences rétrovirales endogènes selon l'invention.

EXEMPLE 6 : Construction d'une souris transgénique contenant tout ou partie du gène de l'envérine.

Une souris transgénique contenant tout ou partie de la séquence HERV-7q (SEQ ID NO:3) est construite afin d'identifier les séquences responsables de la spécificité tissulaire, et pour évaluer le rôle de tout ou partie des motifs rétroviraux endogènes de type HERV-7q, en particulier tout ou partie des motifs peptidiques de l'envérine. La technique de micro-injection utilisée se réfère à la technique classique (Hogan et coll., (1994), Manipulating the mouse embryo, Cold Spring Harbor, Cold Spring Harbor Laboratory Press) ou à ses équivalents. Des formes identiques à la molécule humaine normale de motifs de type HERV-7q, dont l'envérine, ou des formes mutées, délétées, présentant des insertions ou tronquées sont testées afin de déterminer les motifs critiques tant sur le plan normal que pathologique, et plus particulièrement au cours du développement foetal et lors des processus tumoraux.

Bibliographie :

- Benit L. et al., 1997. Cloning of a new murine endogenous retrovirus MuERV-L, with

- strong similarity of the human HERV-L element and with a *gag* coding sequence closely related to the Fv1 restriction gene. J. Virol. 71, 5652-5657.
- Coffin J.M. 1985. Endogenous retrovirus. In: "RNA tumor viruses" (Weiss R.A., Varmus H.E., Teich N.M., and Coffin J.M. eds), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Conrad B., Weissmahr R.N., Boni J., Arcari R., Schupbach J., and Mach B. 1997. A human endogenous retroviral superantigen as candidate autoimmunogene in type 1 diabetes. Cell 90, 303-313.
- Covey S.N. 1986. Amino acid sequence homology in *gag* region of reverse transcribing elements and the coat protein gene of cauliflower mosaic virus. Nucleic Acids Res. 14, 623-633.
- Hertig C., Coupar B.E., Gould A.R., and Boyle D.B. 1997. Field and vaccine strains of fowlpox virus carry integrated sequences from the avian retrovirus, reticuloendotheliosis virus. Virology 235, 367-376.
- Hohenadl C., Leib-Mösch C., Hehlemann R., and Erfle Y. 1996. Biological significance of human endogenous retroviral sequences. J. Acqui. Imm. Def. Synd. Hum. Retrovir. 13, S268-S273.
- Kulkoski J.K., Jones S., Katz R.A., Mack J.P.G., and Skalka A.M. 1992. Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases. Mol. Cell. Biol. 12, 2331-2338.
- La Mantia G. et al, N.A.R., 1991, 19, 7, 1513-1520
- Patience C., Wilkinson D.A., and Weiss R.A. 1997. Our retroviral heritage. Trends Genet. 13, 116-120.
- Pearson W.R. 1994. Using the FASTA program to search protein and DNA sequence databases. Methods Mol. Biol. 24, 307-331.
- Perron H., Garson J.A., Bedin F., Beseme F., Paranhos-Baccala G., Komurian-Pradel F., Mallet F., Tuke P.W., Voisset C., Blond J.L., Lalande B., Seigneurin J.M., Mandrand B. and the Collaborative Research Group on Multiple Sclerosis. 1997. Molecular identification of a novel retrovirus repeatedly isolated from patients with multiple sclerosis. Proc. Natl. Acad. Sci. USA 94, 7583-7588.

- Tönjes R.R. et al., J. AIDS and Hum. Retrovirol, 1996, 13, S261-S267
- Vitelli R., Chiarillo M., Lattero D., Bruni C.B., and Bucci C. 1996. Molecular cloning and expression analysis of the human Rab7 GTP-ase complementary deoxyribonucleic acid. Biochem. Biophys. Res. Commun. 229, 887-890.
- 5 - Weber L.T., Miller M., Jaskolski M., Leis J., Skalka M., and Wlodawer A. 1989. Molecular modeling of the HIV-1 protease and its substrate binding site. Science 243, 928-931.
- Wilkinson D., Mager D.L., and Leong J.A.C. 1994. Endogenous human retroviruses. In: " The Retroviridae " (Levy, J.A. ed), Plenum Press New York, , Vol. 3, 465-535.
- 10 - Xiong Y., and Eickbush, T. 1990. Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J. 9, 3353-3362.

Ainsi que cela ressort de ce qui précède, l'invention ne se limite nullement à ceux de ses modes de mise en œuvre, de réalisation et d'application qui viennent d'être décrits de façon plus explicite ; elle en embrasse au contraire toutes les

15 variantes qui peuvent venir à l'esprit du technicien en la matière, sans s'écarter du cadre, ni de la portée, de la présente invention.

REVENDICATIONS

1°) Fragment d'acide nucléique purifié, caractérisé en ce qu'il comprend tout ou partie d'une séquence codant pour une séquence rétrovirale endogène humaine, qui présente au moins des motifs rétroviraux de type *env*, répondant à la séquence SEQ ID NO:1 ou à une séquence présentant un niveau d'homologie avec ladite séquence SEQ ID NO:1 supérieur ou égal à 80% sur plus de 190 nucléotides ou supérieur ou égal à 70% sur plus de 600 nucléotides pour les domaines de type *env*.

2°) Fragment d'acide nucléique selon la revendication 1, caractérisé en ce qu'il présente à la fois de motifs rétroviraux correspondant à un domaine *env* et répondant à la séquence SEQ ID NO:1 et des motifs rétroviraux correspondant à un domaine *gag* et répondant à la séquence SEQ ID NO:2 ou à une séquence présentant un niveau d'homologie supérieur ou égal à 80% sur plus de 190 nucléotides ou supérieur ou égal à 70% sur plus de 600 nucléotides pour les domaines de type *env* et un niveau d'homologie supérieur ou égal à 90% sur plus de 700 nucléotides ou supérieur ou égal à 70% sur plus de 1200 nucléotides pour les domaines de type *gag*, lesquels motifs ne présentent aucune insertion ou délétion supérieure à 200 nucléotides.

3°) Fragment d'acide nucléique, caractérisé en ce qu'il comprend un segment d'une séquence selon la revendication 1 ou la revendication 2 et notamment les séquences SEQ ID NO:3-22, 28 et 61, les séquences nucléiques complémentaires et les séquences inverses complémentaires des séquences précédentes ainsi que les fragments issus des régions codantes des séquences précédentes correspondant à un cadre glissant supérieur ou égal à 14 nucléotides ou leurs séquences complémentaires.

4°) Transcrits, caractérisés en ce qu'ils sont générés à partir des séquences selon l'une quelconque des revendications 1 à 3.

5°) Réactif de diagnostic pour la détection différentielle de séquences nucléiques endogènes humaines complètes ou partielles, présentant des motifs rétroviraux, sélectionnés parmi les séquences SEQ ID NO:1 et/ou SEQ ID NO:2, caractérisé en ce qu'il est sélectionné dans le groupe constitué par les séquences SEQ ID NO:1-22, 28, 37-57, 59-61 et 121-122, les séquences nucléiques complémentaires et les séquences inverses complémentaires des séquences précédentes, par les fragments nucléotidiques capables de définir ou d'identifier les séquences SEQ ID

NO:1 et/ou SEQ ID NO:2 et toute séquence flanquante ou les chevauchant ainsi que par les fragments issus des régions codantes des séquences SEQ ID NO:1-22 et 61, correspondant à un cadre glissant supérieur ou égal à 14 nucléotides ou leurs séquences complémentaires, éventuellement marquées avec un marqueur approprié.

5 6°) Réactif selon la revendication 5, caractérisé en ce qu'il est choisi dans les régions situées entre les nucléotides 3065 et 4390, les nucléotides 6965 et 9550 ou les nucléotides 2502-2865 de la SEQ ID NO:3.

 7°) Réactif selon la revendication 5, caractérisé en ce qu'il est sélectionné parmi les séquences SEQ ID NO:37-57, 59-60 et 121-122 et en ce qu'il est apte
10 à être utilisé comme amorce.

 8°) Réactif selon la revendication 5, caractérisé en ce qu'il est sélectionné parmi les séquences suivantes :

 - un fragment de 1505 nt amplifié par le couple d'amorces SEQ ID NO:37 et SEQ ID NO:38 (amorces G1F et G1R),

15 - un fragment de 2529 nt amplifié par le couple d'amorces SEQ ID NO:45 et SEQ ID NO:46 (amorces E1F et E1R)

 - un fragment de 182 nucléotides répété deux fois, situé en amont du domaine *gag* aux positions 2502-2611/2613-2865

 et en ce qu'il est apte à être utilisé comme sonde.

20 9°) Réactif selon la revendication 5, caractérisé en ce qu'il est choisi dans le groupe constitué par les fragments codants ou non-codants pour tout ou partie de l'envérine et notamment les fragments comprenant au moins 14 nucléotides et plus particulièrement les fragments codant pour la partie C-terminale de l'envérine, soit à partir de l'acide aminé 291, soit à partir de l'acide aminé 321, à compter du codon
25 codant pour la première méthionine.

 10°) Procédé de détection rapide et différentiel des séquences nucléiques rétrovirales endogènes de type *env* ou *env* et *gag*, de leurs variants normaux ou pathologiques, par hybridation et/ou amplification génique, réalisé à partir d'un échantillon biologique, lequel procédé est caractérisé en ce qu'il comprend :

30 (a) une étape dans laquelle l'on met en contact un échantillon biologique à analyser avec au moins une sonde selon la revendication 5, la revendication 6

ou la revendication 8 et

(b) une étape dans laquelle on détecte par tout moyen approprié le ou les produits résultants de l'interaction séquence nucléotidique-sonde.

11°) Procédé de détection selon la revendication 10, caractérisé en ce qu'il comprend :

* préalablement à l'étape (a) :

. une étape de préparation du tissu ou du liquide biologique concerné,

. une étape d'extraction de l'acide nucléique à détecter, et

10 . au moins un cycle d'amplification génique mis en œuvre à l'aide d'au moins un réactif selon l'une quelconque des revendications 5 à 7 et

* postérieurement à l'étape (b) :

15 . une étape de comparaison des séquences nucléiques obtenues dans ledit échantillon biologique avec les séquences rétrovirales endogènes humaines selon l'une quelconque des revendications 1 à 3, par tout moyen approprié et notamment par séquençage, Southern-blot, coupure de restriction, SSCP ou toute autre méthode permettant d'identifier une insertion ou une délétion ou encore une simple mutation entre les différentes séquences comparées.

12°) Procédé de détection des transcrits selon la revendication 4, caractérisé en ce qu'il comprend :

- le prélèvement des ARN messagers provenant d'échantillons biologiques témoins et d'échantillons analogues prélevés chez des patients et

25 - l'analyse qualitative et/ou quantitative desdits ARNm, par hybridation *in situ*, par dot-blot, Northern-blot, RNase mapping ou RT-PCR, à l'aide d'un réactif de diagnostic selon l'une quelconque des revendications 5 à 9.

13°) Séquences chimères, caractérisées en ce qu'elles sont constituées par un fragment de 17 à 40 nucléotides d'une séquence flanquante sélectionnée dans le groupe constitué par les transcrits et ADNc des séquences génomiques, qui codent pour tout ou partie d'un facteur, dont la fonction, la régulation/dérégulation ou 30 l'altération est associée à l'expression normale ou pathologique ou à la régulation/dérégulation de motifs appartenant à ladite famille HERV-7q, ces séquences

correspondant à des séquences nucléotidiques codant pour des gènes situés dans des régions flanquantes situées en amont et/ou en aval d'une séquence rétrovirale de ladite famille HERV-7q dont l'une des extrémités ne peut se trouver à une distance excédant 120 kb, associée à un motif rétroviral endogène de type HERV-7q comprenant entre
5 17 et 40 nucléotides selon les revendications 1 à 4.

14°) Méthode de détection et/ou d'évaluation d'une sur-expression/sous-expression ou d'une modification d'au moins l'une des séquences ou fragments de séquences rétrovirales endogènes de type HERV-7q et/ou de leurs séquences flanquantes associées, selon l'une quelconque des revendications 1 à 9, caractérisée en
10 ce qu'elle comprend :

- le dépôt sur un support approprié, de l'ADNc provenant de clones, de produits de PCR obtenus à partir d'ADN génomique, de produits de RT-PCR provenant de transcrits ou encore de séquences oligonucléotidiques spécifiques, lesdites séquences d'ADN étant des séquences ou des fragments de séquences rétro-
15 virales endogènes de type HERV-7q et/ou leurs séquences flanquantes, constituées par les transcrits et ADNc des séquences génomiques, qui codent pour tout ou partie d'un facteur, dont la fonction, la régulation/dérégulation ou l'altération est associée à l'expression normale ou pathologique ou à la régulation/dérégulation de motifs appartenant à ladite famille HERV-7q, ces séquences correspondant à des séquences
20 nucléotidiques codant pour des gènes situés dans des régions flanquantes situées en amont et/ou en aval d'une séquence rétrovirale de ladite famille HERV-7q dont l'une des extrémités ne peut se trouver à une distance excédant 120 kb et/ou une séquence chimère selon la revendication 13,

- l'hybridation dudit support avec au moins une sonde marquée de
25 manière adéquate obtenue, par exemple, par rétrotransposition d'un mélange d'ARN provenant de cellules, de tissus ou de liquides biologiques provenant de témoins réputés normaux, de membres de populations ethniques différentes, de patients atteints de pathologies souvent associées à une expression de rétrovirus, comme les processus tumoraux, ou comme les maladies auto-immunes, et

30 - la détection des hybrides formés.

15°) Méthode selon la revendication 14, caractérisée en ce que ledit

transcrit ou ADNc est sélectionné dans le groupe constitué par les séquences SEQ ID NO:62-67 et 119 et leurs fragments correspondant à un cadre glissant supérieur ou égal à 14 nucléotides ou leurs séquences complémentaires.

16°) Méthode selon la revendication 14 ou la revendication 15.
5 caractérisée en ce que ledit support comprend en outre toute séquence rétrovirale endogène ou exogène.

17°) Kit de détection et/ou d'évaluation d'une maladie auto-immune et notamment des neuropathologies à étiologie auto-immune, caractérisé en ce qu'il comprend outre les tampons nécessaires à la mise en œuvre d'un procédé selon l'une
10 quelconque des revendications 14 à 16 :

- des réactifs A de diagnostic selon l'une quelconque des revendications 5 à 9, et

- des réactifs B constitués par les transcrits et ADNc des séquences génomiques, qui codent pour tout ou partie d'un facteur, dont la fonction, la régulation/dérégulation ou l'altération est associée à l'expression normale ou pathologique
15 tion/dérégulation ou l'altération est associée à l'expression normale ou pathologique ou à la régulation/dérégulation de motifs appartenant à ladite famille HERV-7q, ces séquences correspondant à des séquences nucléotidiques codant pour des gènes situés dans des régions flanquantes situées en amont et/ou en aval d'une séquence rétrovirale de ladite famille HERV-7q dont l'une des extrémités ne peut se trouver à une distance
20 excédant 120 kb,

lesquels réactifs sont de préférence fixés sur un support approprié.

18°) Kit selon la revendication 17, caractérisé en ce que lesdits réactifs B sont sélectionnés dans le groupe constitué par les séquences SEQ ID NO:62-67 et 119 et leurs fragments correspondant à un cadre glissant supérieur ou
25 égal à 14 nucléotides ou leurs séquences complémentaires.

19°) Produits de traduction, caractérisés en ce qu'ils sont codés par une séquence nucléotidique selon l'une quelconque des revendications 1 à 4.

20°) Peptide, caractérisé en ce qu'il est susceptible d'être exprimé à l'aide d'une séquence nucléotidique sélectionnée dans le groupe constitué par les
30 séquences SEQ ID NO:1-22, 28 et 61 selon l'une quelconque des revendications 1 à 4.

21°) Peptide selon la revendication 20, caractérisé en ce qu'il

englobe les peptides dérivés comprenant entre 5 et 540 aminoacides et notamment un fragment de 538 aminoacides, commençant à la première méthionine de la séquence SEQ ID NO:26 (envérine).

- 22°) Peptide selon la revendication 20 ou la revendication 21,
5 caractérisé en ce qu'il est sélectionné dans le groupe constitué par :
- . les séquences SEQ ID NO:23-36 ;
 - . la séquence SEQ ID NO:58 ;
 - . un fragment C-terminal de la séquence SEQ ID NO:26, soit à partir de l'acide aminé 291, soit à partir de l'acide aminé 321, à compter de la première
- 10 méthionine de la séquence SEQ ID NO:26 ;
- un peptide de type CKS-17/CKS-25 présent dans l'une des séquences SEQ ID NO:23-36 ou 58 ; et
 - les peptides présentant une affinité avec un des haplotypes du système HLA de classe I ou de classe II et notamment les fragments 399-471, 244-271
- 15 de l'envérine, ainsi que les peptides de séquence SEQ ID NO:68-118, conformément au Tableau I.

- 23°) Peptide selon l'une quelconque des revendications 20 à 22, caractérisé en ce qu'il est obtenu à partir des séquences nucléiques selon l'une quelconque des revendications 1 à 4, dans lesquelles au moins un codon non-sens peut
- 20 être remplacé par un codon codant pour l'un des aminoacides suivants : Phe (F), Leu (L), Ser (S), Tyr (Y), Cys (C), Trp (W), Gln (Q), Arg (R), Lys (K), Glu (E) ou Gly (G).

- 24°) Compositions immunogènes ou vaccinales, pour la protection contre les maladies auto-immunes, notamment chez les sujets à risque, caractérisée en
- 25 ce qu'elle comprend au moins un peptide comprenant au moins un motif de type CKS et/ou au moins un motif sélectionné dans le groupe constitué par les peptides présentant une affinité avec un des haplotypes du système HLA de classe I ou de classe II et au moins un véhicule pharmaceutiquement acceptable.

- 25°) Composition selon la revendication 24, caractérisée en ce que
- 30 ledit peptide présentant une affinité avec un des haplotypes du système HLA de classe I ou de classe II, est sélectionnée dans le groupe constitué par les peptides tels que

définis dans le Tableau I.

26°) Composition selon la revendication 24 ou la revendication 25, caractérisée en ce que ledit peptide présente la séquence SEQ ID NO:120.

27°) Anticorps, caractérisé en ce qu'il est dirigé contre l'un ou
5 plusieurs des peptides selon l'une quelconque des revendications 20 à 23.

28°) Composition pharmaceutique, caractérisée en ce qu'elle comprend des anticorps neutralisants produits à partir des peptides du Tableau I (SEQ ID NO:68-118) et leurs homologues.

29°) Procédé de dépistage immunologique différentiel de séquences
10 rétrovirales endogènes humaines de la famille HERV-7q normales ou pathologiques, caractérisé en ce qu'il comprend la mise en contact d'un échantillon biologique avec un anticorps selon la revendication 27, la lecture du résultat étant révélée par un moyen approprié, notamment EIA, ELISA, RIA, fluorescence.

30°) Procédé d'identification et de détection de motifs rétroviraux
15 endogènes, anormalement exprimés dans le cadre de pathologies associées au cancer, ou de neuropathologies en particulier autoimmunes, au premier rang desquelles la sclérose en plaques, caractérisé en ce qu'il comprend l'analyse comparée des séquences extraites d'un échantillon biologique avec les séquences selon l'une quelconque des revendications 19 à 23.

20 31°) Application des séquences selon l'une quelconque des revendications 1 à 9, 13, 14 ou 19 à 23 au diagnostic, au pronostic, à l'évaluation de la susceptibilité génétique, à toutes maladies humaines induites, innées ou acquises en particulier celles à composantes cancéreuses, autoimmunes et/ou à incidence neurologique, comme la sclérose en plaques, les syndromes associés et les maladies neuro-
25 dégénératives où intervient tout ou partie des séquences selon l'une quelconque des revendications 1 à 5 et des formes endogènes ou exogènes apparentées.

32°) Séquences nucléiques hybrides, caractérisées en ce qu'elles comprennent des séquences ou motifs selon l'une quelconque des revendications 1 à 9, combinés avec des séquences ou motifs d'origine endogène ou d'origine ou induits
30 de manière exogène.

33°) Vecteur recombinant de clonage ou d'expression, caractérisé en

ce qu'il comprend une séquence nucléique selon l'une quelconque des revendications 1 à 4.

34°) Composition immunogène ou vaccinale, caractérisée en ce qu'elle comprend un vecteur incluant au moins une séquence nucléique codant un peptide tel que défini dans le Tableau I, éventuellement associée à une séquence codant un motif de type CKS-17.

35°) Vecteur de thérapie génique, caractérisé en ce qu'il comprend tout ou partie des séquences nucléiques rétrovirales endogènes de type HERV-7q selon l'une quelconque des revendications 1 à 4.

36°) Vecteur selon la revendication 35, caractérisé en ce que lesdites séquences sont sélectionnées dans le groupe constitué par les séquences SEQ ID NO:2, 20 et 21.

37°) Animaux transgéniques, caractérisés en ce qu'ils comprennent tout ou partie d'une séquence de type HERV-7q (SEQ ID NO:1-22 , 28 et 61).

1/64

CCCTGGGGCGGGCTTCCTTTCTGGGATGAGGGCAAAACGCTGGAGATACAGCAATTATCTTGCAACTGAG	71	
AGACAGGACTAGCTGGATTTCCTAGGCGGACTAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACCACGTCCAC	143	
CTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAATGCTAATTAGGCA	215	
AAGACAGGAGGTAAAGAAATAGCCAAATCATCTATTGCTTGAGAGCACAGCAGGAGGGACAACAATCGGGATA	287	
TAAACCCAGGCATTGAGCTGGCAACAGCAGCCCCCTTTGGGTCCTTTCCCTTTGTATGGGAGCTGTTTTTC	359	<u>région</u>
ATGCTATTTCACTCTATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCT	431	<u>répétée</u>
TTTGCTCACCGTCCACCACTGCTGTTTGGCCACCACCGCAGACCTGCCGTGACTCCCATCCCTCTGGATCCT	503	<u>R1</u>
GCAGGGTGTCCGCTGTGCTCTGATCCAGCGAGGCGCCCATGCGCTCCCAATTGGGCTAAAGGCTTGCCA	575	
TTGTTCCTGACGGCTAAGTGCCTGGGTTTGTCTAATTGAGCTGAACACTAGTCACTGGGTTCATGGTTC	647	
TCTTCTGTGACCCACGGCTTCTAATAGAATAAACAATTACCACTTACCACATGGCCCAAGATTCCATTCTTGGAAT	719	
CCGTGAGGCCAAGAACTCCAGGTCAGAGAATACGAGGCTTGCCACCATCTTGGAAGCGGCTGCTACCATCT	791	
TGGAAGTGGTTCACCACCATCTTGGGAGCTCTGTGAGCAAGGACCCCCGGTAACATTTTGGAACCCACGAA	863	
CGGACATCCAAAGTGGTGAGTAATATTGGACCACCTTCACTTGTCTATTCTGTCTTCTCTAGAAATG	935	
GAGGAAAATCCGGGCACCTTGTGGCCAGTTAAACAGATTAGTGTGGCCACCGGACTTAAGACTCAGGTGT	1007	
GAGGCTATCTGGGAAGGGCTTTCTAACAACCCCCAACCTTCTGGGTGGGGACTTGGTTGCTCAAGCC	1079	
AGCTTCCACTTTCAGTTTTCTTGGGAAGCCGAGGGCCGACTAGAGGCAGAAAGCTGTCTCTGAACCTCC	1151	
GGCAGTAGCCGGTTGAGATCATGGTGTAGCCAGAAGTCTCAACAGTCGCCCATGCATGCACCCCTATCTTTC	1223	
CTTCTGACCCATACCTCCTGGGTCCCAACCACTTTCTTCAAAGTGTAGCCCCAAAATCTCTTACCTC	1295	
TGAATATACTTCTCTGATCCCTGCCCTCCTAGGTACTATTGGTTCAGACTTCCATTCTCTAGCAAGTTGT	1367	
ATCTCCAAAGGATCTAAGGAAGCTCTGCGCTGCGTCTTAGGCACCTAGGCTATAACCCAGGGAGTCTTAT	1439	
CCCTGGTGTCTCCCTCCCACTTTAGGCATACAGCTCTTGACATGGGCAGTTATGTAGGACCCACTCCCCACCAC	1511	
CCTTGCCAGGGCCCCAAGTTTGTAAATGGCTGAGGAAAAGAGAGACAGAGAGAGAGAGAAATGGAGGA	1583	
GAAAGAGAGAGAGACAGAGAGAGAGAGAGACAGTGAAGAGACAGAAAGAGAGAGAGACAAAGAGGAGAG	1655	
AGAGAGAGTCAAAGAGAGAAAGAAAGAGAAATAGTAAACAGTGTGCCCTATTCTTTAAAGCCA	1727	
GGGTAAATTTAAACCTGTACTTGTATAATTGAAGGCTTCTCTGTGACCTATAGCACTCCAATCCACTTTG	1799	
TGGTCAGTGTAATAAGAGCATAGGCCGAAAGCACTGAGGCCATTGACAACCCGTAGCTTCCCTATCAAAAA	1871	
TCCTTAACCCAGTAACCCGCAGATGGACCAATGCATTAGTCGGTAGCGCACTGCTTGTCTAAAGTAGA	1943	
AAAGTAACTTTTAGAGGAAACCTCATTGTGAGCACACCTCACCTGTCAGAATTATCTAATAAAAAAGCA	2015	
AAAAGGTAGCTTACTAATCAAAAATCTTAAAGTATGGGCTATTCTGTAGAAAAAGGTAATGTAATCCA	2087	
ACCACTGATAAATCCCTTAACCCAGCAGATTTCTTAACGGGATTTAAATCTTAATTACCATACAAAGGTCCG	2159	
ACCAGACCTTAGCGGAACTCCCTTCAGGACAGGACGATAGTGGTTCTCCAGGTGATTGAGGAAAAAAC	2231	
CACAATGGGTATTCAAGTAATTGATACGGGACTCTTGTGGAAGCAGAGTTAGAAAAATGCCTAATAACTGG	2303	
TCTCTCAAACGTGTGAGCTGTTTGCCTCAGCCAAGCCTTAAAGTACTTACAGAATCAAAAGACTATCTCA	2375	
ATCCTGATTTCAAAGGTTAGCTACACCTCTCTGTAATGCATTGTGATAAGAACTTGTATTGGAATGCAT	2447	
CTTGATGGGGCAGCTGGGTTGTTATAAAATAGGAACCCAGCCAGCTCTAGGACTCACCCCTGAGCGCAAAG	2519	
GCAATGTTGGGCATGCTGGTAAAGGACCACTAGAATCCAGCAGCCAGACCCCTTTCTTTGTGGTCAAGAAA	2591	<u>régions</u>
GGCGGGAAAAGGGGTGAGGACTGTCTACATCGGTAAAGCATACTAATCCGATAAACAGAGGTCCATGGGTGG	2663	<u>répétées</u>
TTACGCACCTTGAAAGGAACCTCACCCCTGAGCACAAAGGCAATGTTGGGCACGCTGGTAAAGGACCACTAG	2735	<u>en tandem</u>
AATCCAGCAGCCTGGACCCCTTTCTTTTGGTCAAGAGAGGGCAGGAAAACAGGTGCAGGACTGCAACATCAG	2807	<u>R2</u>
TGAGCATACTAATTGATAAGCAGAGGTCCATGGGTGGTGTGACCCCTGGAAGAAATAAGCATTAGGACC	2879	
ATAGAGGACACTCCAGGACTAAAGCTCATCGGAAAATGACTAGGGTTGCTGGCATCCCTATGTTCTTTTTTC	2951	
AGATGGGAACCGTTCCCGCAAGACAAAAACGCCCTAAGACGTATTCTGGGAATTGGGACCAATTTGACC	3023	
CTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCACTGCCGCTTGGCCTCTGAGGGAAGTATAAAT	3095	
TATAACACCATCTTACAGCTAGACCTCTTTGTAGAAAAGGCAATGGAGTGAAGTGCCATAAGTACAACT	3167	
TTCTTTTCATTAGAGACAACCTACAATTATGTAAAGTGTGATTTATGCCCTACAGGAAGCCTTCAGAGT	3239	
CTACCTCCCTATCCCGACATCCCGACTCCTTCCCCAACTAATAAGGACCCCCCTTCAACCCAAATGGTCCA	3311	
AAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCAAATATCCCCAATTATGACCCCTCCAAGC	3383	
AGTGGGAGGAAGAGAATTCCGGCCAGCCAGAGTGATGTGCCTTTTCTCTCCAGACTTAAAGCAAAATAAA	3455	
AACAGACTTAGGTAAATTTCTCAGATAACCTGATGGCTATTGATGTTTACAGGGTTAGGACAATTCTT	3527	
TGATCTGACATGGAGAGATATAATGTCAGTCTAAATCAGACACTAACCCCAATGAGAGAAGTGCCACCAT	3599	
AACTGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCATATGATAGGATGACAACAGAGGAAAG	3671	<u>domaine</u>
AGAATGATTCCCCACAGGCCAGGCAGTTCAGTCTAGACCTCATTTGGGACACAGAATCAGAACATGG	3743	<u>gag</u>
AGATTGGTGTGCAGACATTGCTAACTTGTGTCTAGAAGGACTAAGGAAAACTAGGAAGAAGTCTATGAA	3815	
TTACTCAATGATGTCCACCATAACACAGGGAAGGGAAGAAATCCTACTGCTTTCTGGAGAGACTAAGGGA	3887	
GGCATTTGAGGAAGCGTGCTCTCTGTCACTGACTCTTCTGAAGGCCAATACTTAAAGCGTAAGTTTAT	3959	
CACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAGTCTGCCGTAGGCCCGGAGCAAACTTAGAAAC	4031	
CCTATTGAACTTGGCAACCTCGGTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAACAGGACAAACGGGA	4103	
TTAAAAAAAAGGCCACCGCTTTAGTCATGACCCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGAAAA	4175	
GCTGGGCAAAATGAATGCTAATAGGGCTTGCTTCCAGTGCCTTACAAGGACACTTTAAAAAAGATTGTC	4247	
CAAGTAGAAGTAAGCGCCCTCGTCCATGCCCTTATTTCAGGGAATCACTGGAAGGCCACTGCCCCA	4319	
GGGGACAAGGTCTCTGAGTCAGAAGCCACTAACAGATGATCCAGCAGCAGGACTGAGGGTGCTGGGGC	4391	
AAGCGCCATCCCATGCCATCACCCCTCACAGAGCCCTGGGTATGCTTGACCATGAGGGCCAGGAGTTGTCT	4463	
CCTGGACACTGGTGCGGCTTCTTAGTCTTACTCTTCTGCCGACAACTGTCTCCAGATCTGTCACTAT	4535	
CTGAGGGGCTCTAAGACGGGCAGTCACTAGATACTTCTCCAGCCACTAAGTTATGACTGGGGAGCTTTAT	4607	
TCTTTTCACTAGCTTTTCTAATTATGCTTGAAGCCCCACTACCTTGTTAGGGAGAGACATTCTAGCAAAAG	4679	
CAGGGGCCATTATACACCTGAACATAGGAGAAGGAACACCCGTTTGTGTCCCCTGCTTGAGGAAGGAATTA	4751	
ATCCTGAAGTCTGGGCAACAGAGAGGACAATATGGACGAGCAAGAAATGCCGCTCTGTTCAAGTTAACTAA	4823	
AGGATTCCACCTCCTTTCCCTACCAAGGCGAGTACCCCTCAGACCCAAGGCCCAACAAGGACTCCAAAAGA	4895	
TTGTTAAGGACCTAAAAGCCCAAGGCCCTAGTAAACCATGCGATAACCCCTGCAGTACTCCAATTTAGGAG	4967	
TACAGAAACCAACAGACAGTGGAGGTTAGTGCAAGATCTCAGGATTATCAATGAGGCTGTTGTTCTCTAT	5039	<u>domaine</u>
AGCCAGCTGTACCTAGCCCTTATACTCTGCTTCCCAATACCAGAGGAAGCAGAGTGGTTTACAGTCTTG	5111	<u>pol</u>
ACCTTCAGGATGCCTTCTTCTGCATCCCTGTACATCTGACTCTCAATTCTTGTTCCTTTGAGAGATACTT	5183	

FIGURE 1.1

2/64

CAAACCCAACTCTCAACTCACCTGGACTATTTACCCCAAGGGTTTCAGGGATAGTCCCATCTATTGGCC
 AGGCATTAGCCCAAGACTTGAGCCAACTCTCATACCTGGACACTTGTCTTCGGTAGGTGGATGATTACTT
 TTGGCCGCCCATTCAGAAACCTTGTGCCATCAAGCCACCCAAAGCGCTCTTCAATTTCTCGCTACCTGTGGC
 TACATGGTTTCCAAACCAAGGCTCAACTCTGCTCAGCAGGTTACTTAGGGCTAAAATTATCCAAAGGCA
 CCAGGGCCCTCAGTGAGGAACACATCCAGCCTATACTGGCTTATCTCATCCCAAAACCCTAAAGCAACTAA
 GGGGATTCTCTTGGCGTAATAGGTTTCTGCCGAAAATGGATTCCAGGTATGGCGAAATAGCCAGGTCAATTA
 ATACACTAATTAAGGAACTCAGAAAGCCAATACCCATTTAGTAAGATGGACAACCTGAAGTAGAAGTGGCTT
 TCCAGGCCCTAACCCAAAGCCCAAGCTGTTAAGTTTGCCAAACAGGGCAAGACTTTTCTCATATGTACAGAAA
 AAACAGGAATAGCTCTAGGAGTCTTACACAGATCCGAGGGATGAGCTTGCAACCTGTGGCATACTGACTA
 AGGAAATTGATGTAGTGGCAAGGGTTGACCTCATTGTTTACGGGTAGTGGTGGCAGTAGCAGTCTTAGTAT
 CTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCA
 CTGCTAAAGGAGACTTGTGGCTGTGACACAACCTGTTTACTTAAATGTCAGGCTCTATTACTTGAAGGGCCAG
 TGCTGCGACTGTGCACTTGTGCACTCTTAACCCAGCCACATTTCTTCCAGACAATGAAGAAAAGATAAAAC
 ATAAGTGTCAACAAGTAATTTCTCAACCTATGCCACTCGAGGGGACCTTTTAGAGGTTCCCTTGACTGATC
 CGAGCTCAACTGTATCTAGTGAAGTTCCCTTGTAGGAAAAGGACTTCGAAAAGTGGGGTATGCAGTGG
 TCAGTGATAATGGAATACTTGAAGTAATCCCTCACTCCAGGAAGTGTGCTCAGTAGCAGAACTAATAG
 CCTCACTTGGGCACTAGAATTAGGAGAAGAAAAGGGCAATATATATACAGACTCTAAATATGCTTACC
 TAGTCTCCATGCCATGCAGCAATATGGAAGAAAGGGAATCTTAACCTTCTGAGAGAACCTATCAAAAC
 ATCAGGAAGCCATTAGGAAATTTATTGGCTGTACAGAAACCTAAAGAGGTGGCAGTCTTACTGCGGGG
 GTCATCAGAAAGGAAAGGAAAGGAAATAGAAGAGAAGTGCACAGCAGATATTGAAGCCAAAAGAGCTGCAA
 GGCAGGACCTCCATTAGAAATGCTTATAAAACAACCCCTAGTATAGGGTAATCCCTCCGGGAAACCAAGC
 CCCAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCAGAGGACAGTTTTCTCCCTCCGGGACGGCTAGCC
 ACTGAAGAAAGGAAAATACTTTTGGCTGCAACTATCCAATGGAATTAAGTTAAACCCCTTCACTCAAACTTT
 CACTTAGGCATCGATAGCACCATCAGATGGCCAAATCATATTACTGAGCAGGCGCTTTTCAAACTATC
 AAGCAGATAGTCAGGGCTGTGAAGTGTGCGAGAGAAATATCCCTGCTTATCGCCAAAGCTCTTCAGGA
 GAACAAAGAACAGGCCATTACCTTGGAGAGACTGGCAACTGATTTTACCCACAAGCCAAACCTCAGGGAT
 TTCAGTATCTACTAGTCTGGGTAGATACTTTCACGGGTTGGGCAAGGCGCTTCCCTGTAGGACAGAAAAGG
 CCCAAGAGGTAATAAGGCACTAGTTCATGAAATAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTG
 ACAATAGCCCTGCTTTCAGGCCACAGTAACCCAGGGAGTATCCAGGCGTTAGGTATACGATATCACTTAC
 ACTGCGCTGAAGGCCACAGTCTCAGGGAAGGTGAGAGAAATGAATGAAACACTCAAAGGACATCTAAAAA
 AGCAAAACCCAGGAAACCCACCTCACATGGCTGCTGTGTGCTTATAGCCTTAAAAAGAACTGCAACTTC
 CCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTATAACCAATGACCTTGTGCTTG
 ACCCAAGACAGCCCACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTCTTAAAAACATTACA
 AGGAACCTATCCCTGAGAAAGGGGAAAGAACTATTCACCCCTGTGACATGGTATTAGTCAAGTCCCTTCC
 TCTAATTCCTCCATCCCTAGATACATCTGGGAAGGACCTACCCAGTCAATTTATCTACCCCAACTGCGGT
 TAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAATCTGGTACTGCCAAAGGAACCTGA
 AAATCCAGGAGACAACGCTAGCTATTCTGTGAACCTCTAGAGGATTGCGCCTGCTCTTCAAAACAACACC
 AGGAGGAAAGTAATCAATCAATAATCCCATGGCCCTCCCTTATCATATTTTCTCTTACTGTTCTTTT
 ACCCTCTTTCACTCTCACTGCACCCCTCCATGCGCTGTATGACCAAGTCTCCCTTACCAAGAGTTTCT
 ATGGAGAATGCAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTCTAAGGGAACCCCACTT
 CACTGCCCACACCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGATGCATGCAAAATACTCATT
 TTGGACAGGAAAATGATTAATCCTAGTTGCTCTGGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCA
 CCAAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAACATGTAAAAGAAGTAATCTC
 CCAACTCACCCGGGTACATGGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAACCCCT
 CCGTACCCATACTCGCTTGAAGCTATTTAATACCACCTCACTGGGCTCCATGAGGTCTCGGCCAAAA
 CCCTACTAAGTGTGGATATGCTCCCTCCCTGAACTTCAGGCAATATGTTTCAATCCCTGTACCTGAACATG
 GAACAACTTCAGCAGAGAAATAAACACCACTTCCGTTTATAGGAGCTCTGTTTCCAATCTGGAATAAC
 CCATACCTCAAACTCACCTGTGTAATAATTTAGCAATACTACATACAAACCACTCCCAATGCATCAGGTG
 GGTAACTCTCCACACAAATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCG
 TTGTTTGAATGGCTCTTCAGAACTATGTGCTTCTCTCATTTAGTGGCCCTATGACCATCTACACTGA
 ACAAGATTTATACAGTTATGTATATCTAAGCCCCGCAACAAAGAGTACCCATTCTCTTTTGTATAGG
 AGCAGGAGTGTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAACT
 ATCTCAAGAACTAAATGGGGCATGGAAAGGCTGCGCCGACTCCCTGGTCACTTGCAGATCAACTTAACTC
 CCTAGCAGCAGTACTCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGGTGAAAGAGGGGGAACCTGTTT
 ATTTTATAGGGGAAGATGCTGTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCG
 AGATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCCTGGGGCTCTCAGCCAATGGATGCC
 CTGGATTCTCCCTTCTTAGGACCTCTAGCAGCTATAATTTGCTACTCTCTTTGGACCTGTATCTTAA
 CCTCTTGTAACTTTGTCTCTTCCAGAACTGAAGCTGTAATACTACAATGGAGCCCAAGATGCAGTCCAA
 GACTAAGATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCAAAGGCAC
 CCCTCTGAGGAATCTAGCTGCACAACTCTACTACGCCCAATTACAGAGGAAGCAGTTAGAGCGGTC
 TCGGCCAACCTCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGACTGAGAGACAGGACTAGCTGGAT
 TTCTTAGGCTGACTAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTAAACACGGGGCTTG
 CAACCTAGCTCACACCTGACCAATCAGAGAGCTCACTAAATGCTAATTAGGCAAGACAGGAGGTAAGAA
 ATAGCCAATCATCTATTGCTGAGAGCAGCAGGAGGACAAATGATCGGGATATAAACCCCAAGTCTTCGAG
 CCGGCAACGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCATCTAT
 TAAATCTTGAACCTGCATCTTCTGGTCCATGTTTCTTACGGCTTGAGCTGAGCTTTTCGCTCGCCATCCACC
 TAAATCTTGAACCTGCATCTTCTGGTCCATGTTTCTTACGGCTTGAGCTGAGCTTTTCGCTCGCCATCCACC
 ACTGCTGTTTGGCGCCACCGCAGACCCGCGCTGACTCCCTCTGGATCATGAGGGTGTCCGCTGTG
 CTCCTGATCCAGCGAGGCACCCATTGCGCTCCCAATCGGGCTAAAGGCTTGCATTGTTCTGTGATGCCA
 AGTGCCTGGGTTTCATCTAATTGAGCTGAACACTAGTCACTGGGTTCATGTTCTCTTGTGACCCACAG
 CTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGGTTCCATTCCTGAATCCATAAGGCCAAGAACCC
 CAGGTGAGAAACAGAGGCTTGCACCATCTTGGGAGCTCTGTGAGCAAGGACCCCAAGTAACACAACCA
 TGAGGGTGCAAAATGCATGGGCCACTAATGGTAGAGCAAGAAAAGAGGGCCCTGGTTCCTCGAAGGCATC
 AGTGAGCTGAAATGCCTGCCCTGGATGTCTATTCCTAGGTGTTTTCTGCCTGAAGCAGATTAACCCCTT
 GTTCACTTCTCAAGTAGGGCTTCTATTACAGCCCAATCAATCCCAACCCAGATGACAT

5255

5327

5399

5471

5543

5615

5687

5759

5831

5903

5975

6047

6119

6191

6263

6335

6407

6479

6551

6623

6695

6767

6839

6911

6983

7055

7127

7199

7271

7343

7415

7487

7559

7631

7703

7775

7847

7919

7991

8063

8135

8207

8279

8351

8423

8495

8567

8639

8711

8783

8855

8927

8999

9071

9143

9215

9287

9359

9431

9503

9575

9647

9719

9791

9863

9935

10007

10079

10151

10223

10295

10367

10439

10500

domaine

env

 région
 répétée
 R1

FIGURE 1.2

3/64

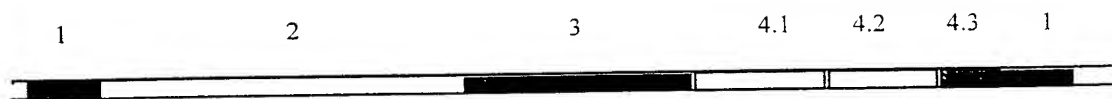


FIGURE 2

[illegible]

FIGURE 3

5/64

IPMALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFWLRMQRPGNIDAPSYRSLSKG
 TPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGVTVWCWYFTQTGMSDGG
 GVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRTHTRLVSLFNNTLTGLHEV
 SAQNPTNCWICLPLNFRPYVSI PVPEQWNNFSTEINTTSVLVGPLVSNLEITHTSNLTCV
 KFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFVCGTSAYRCLNGSSSESMCFLSFLVPPMT
 IYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALGTGIGGITTSTQFYKLSQELNGDM
 ERVADSLVTLQDQLNSLA~~AVVLQNR~~RALDLLTAERGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIQRRAEELR
 NTGPWGLLSQWMPWILPFLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPAS
 RSDVNDIKGTPPEEISAAQPLLRPNSAGSS

FIGURE 4

- 1) NSLA~~AVVLQNR~~RALDLLTAESGGTFLFLEEK
- 2) NSLA~~AVVLQNR~~RALDLLTAERGGTCLFLGEEC
- 3) DSLAAVTLQNHQGLDLLTA~~EKGGLCYFLGEDC~~
- 4) DSLAAVTLQNHQGLDLLIAEKGGGLCTFLGEEC
- 5) DSLAAVTLQNCRGLDLLTA~~EKGGHYTF~~LGEEC
- 6) LQNRRLDLLFLKEGGLC
- 7) DSLAKVVLQNRRLDLLTAEQGGICLALQEK

FIGURE 5

TSFVEKANGVKCHKYKLSFHKETTHNYVKS~~VIYALQEA~~FRVYLPILPASPTPSPTNKDPPSTQM~~VQKEID~~KRVNSEPKS
 ANIPQLXPLQAVGGREFGPARVHVPFSLPDLKQIKTDLGKFSNPDGYIDVLQGLGQFFDLTWRDIMSLLNQTLTPNER
 SATITAAXEFGDLWYLSQVND~~MTTEEREXFPTGQ~~QAVPSLDPHWDTESEHGDWCCRHL~~TCVLEGLR~~KTRKKSMNYSM
 MSTITQGREENPTAFLERLREALRK~~ASLSPDS~~SEGQILKRKFITQSAADIRKKLQKSAVGPEQNLETLLNLATSVFY
 NRDQEEQAEQDKRDXKKGHRFSDPQASGLWRLWKREKLGKLNAXXGLLPVRSTRTLXKRLSKXKXAPSSMPLISRES
 LEGPLPQGTKVLXVRSHXPD/SSSRT

FIGURE 6

FIGURE 7

7/64

```

01/          TAAATCCCCATGGCCCTCCCTTATCATATTTTTTCT
02/          TAAATCCCC-TGGCCCTCCCTTATCATATTTTTTCT
03/          TAAATCCCCATGGCCCTCCCTTATCATATTTTTTCT
04/          TAGATCCTCATGGCCCTCC-TTGTCTATATTTTTTTT

01/CTTTACTGTTCTTTTA-CCCTCTTTCACCTCTCACTGCACCCCCCTCCATGCCGCTGTATGACC
02/CTTTACTGTTCTCTTACCCCCCTTTCACCTCTCACTGCACCCCCGTCCATGCCACTGCACCCCC
03/CTTTACTGTTCTCTTA-CCCCCTTCTCTCTCACTGCACCCCCCTCCATGCTGCTGTACAACC
04/CTTTACTGTTCTCTTA-CCCCCTTTCACCTCTCACTGAACCCCCCTCCATGCCACTGTACTACC

01/AGT-----AGCTCCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGT
02/GTCCATGCCCGTCTCATGCCAGTAGCTCCCCCTTAGCAAGAGTTTCTATGGAGAATGCAGCGT
03/AGC-----AGCTCCCCCTTACCAAGAGTTTCTATGAAGAATGCAGCGT
04/AGT-----AGCTCCCCATTACCAAGAGCTTCTATGGACAATGCAGCGT

01/CCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGC
02/CCCGGAAATATTGATGCCCCATTGTATAGGAGTTTATCTAAGGGAACCCCCACCTTCACTGC
03/CCCGGAAATATTGATGCCCCATCAAATAGGAGTTTACCTAAAGGAAACTCCACCTTCACTGC
04/CCTGGAAATATTGATGACCCATCGTATAGGAGTTTCTAAGGGAACCCCCATTTCACCAC

01/CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCATGCAAATACTC
02/CCACACCCATATGCCCCACAACTGCTATAACTCTGCCACTCTTTGCATGCATGCAAATACTC
03/CCACACCCATATGCCCCACAACTGCTATAACTCTGCCACTCTTTGCATGCATGCAAATACTC
04/CCACACCTATATGACCC-----

01/ATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCTGGAGGACTTGGAGTCACTGTCTGT
02/ATTATTGGACAGGAAAAACGATTAATCCAGTTGTCTGGAGGACTTGGAG-----
03/ATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCTGGAGGACTTGGAGCCACTGTCTGT
04/------

01/TGGACTTACTTCACCCAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGA
02/--GACTCACTTCACTCATACCGATATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAACAGA
03/CGGACTTACTTCACCCATACTGGTATGTCTGAGGGGGGTGGAGTTCAAGATCAGGCAAGAGA
04/------

01/AAAACATGTAAAAGAAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCCTACA
02/AAAACACATAAAGGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCCCCTACA
03/AAAACATGTAAAGGAAGTAACCTCCCAACTGACCCGGGTACATAGCACCCCTAGCCCCCTACA
04/------

01/AAGGACTAGATCTCTCAAACTACATGAAACCCCTCCGTACCCATACTCGCCTGGTAAGCCTA
02/AAGGACTAGATCTCTCAAACTACATGAAACCCCTCCATAACCCATACTGGCCTGGTAAGCCTA
03/AAGGACTAGATCTCTTAAACTACATGAAACCCCTCCATAACCCATACTTGCCTGGTAAGCCTA
04/------

01/TTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTTGGAT
02/TTTAATACCACCCTGACTGGGCTCCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTTGGAT
03/TTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGTCCAAACCCCTACTAACTGTTGGTT
04/------

01/ATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACCT
02/GTGCCTCCCCCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACAATGGAACAACCT
03/GTGCCTCCCCCTGTATTTAGGCCATGCATTTCAATCCCTGTACCTGAACAATGGAACAACCT
04/------TGCACCTCAGGCCATACATTTCAATCCCTGTA-----

```

FIGURE 8.1

8/64

```

01/TCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATA
02/TCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATA
03/ACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATA
-----

01/ACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCA
02/ACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACAGCCAACCTCCCA
03/ACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACAACCAACTCCCA
04/-----

01/ATGCATCAGGTGGGTAACCTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTG
02/ATGCATCAGGTGGGTAACCTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTG
03/ATGCATCAGGTGGGTAACCTCCTCCCACACGAATAGTCTGCCTACCCTCAGGAATATTTTTTG
04/-----

01/TCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCA
02/TCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTCA
03/TCTGTGGTACCTTAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCA
04/-----

01/TTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAA
02/TTCTTAGTGCCCCCTATGCCCATCTACACTGAACAAGATTTATACAATCATGTCATACCTAA
03/TTCTTAGTGCCCCCTATGACCATTTACACTGAACAAGATTTATACAATTATGTTGTACCTAA
04/-----

01/GCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTATAGGAGCAGGAGTGCTAGGTGCAC
02/GCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTATAGGAGCAGGAGTGCTAGGCGGAG
03/GCCCCACAACAAAAGAGTACTCATTTCTTCCTTTTGTATCGGAGCAGGAGTGCTAGGTGGAC
04/-----

01/TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAA
02/TAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAACTGTCTCAAGAA
03/TAGGTCTCTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAA
04/-----

01/CTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTC
02/CTTAAAGGTGACATGGAATGGGTCGCTGATACCCTGGTCACCTTGCAAGATCAACTTAACTC
03/CTCAATGGTGACATGGAATGGGTTGCCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTC
04/-----

01/CCTAGCAGCAGTAGTCCTTCAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGGG
02/CCTAGCAGCAGTAGTCCTTCAAATCGAAGAGCTTTAGACTTGCTAACCGCGGAAAGCGGGG
03/CCTAGCATCAGTAGTCCTTCAAATCGAAGAGCTTTAGACTTGCTAACCTCTGAAAGAGGGG
04/-----

01/GAACCTGTTTTATTTTTAGGGGAAGAATGCTGTTATTATGTT-----
02/GAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTT-----
03/GAAGCTGTTTTATTTTTAGGGGAAGAATGTTGTTATTATGTTATTTTAGCGGAAGAATGTTGT
04/-----

01/-----AATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACA
02/-----AATCAATCCGGAATCATCACCAGAAAGTTAAAGAAATTCAGGTCTGAATATA
03/TATTATGTTAATCAATCCTGAATTGTCACAGAGAAAGTTGAAGAAATTCGAGATTGAATACA
04/-----

01/ACGTAGAGCAGAGGAGCTTCGAAA-CACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCT
02/ACGTAGAGCAAAGGAGCTGCAAAA-CACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCT
03/ACGTAGAACAGAGGAGCTTCAAAAACACCAGACCCTGGGGCCTCCTCAGCCAATGGATGCCCT
04/-----

```

FIGURE 8.2

9/64

```

01/GGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTA
02/GGATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTTGGACCCTGTA
03/GGATTCTCCCCTTCTTAGGATCTCTAGCAGCTCTAATATTGATACCTCCTCTTTGGACCCTGTA
04/-----

01/TCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAACTA-----
02/TCTTTAACCTCCTTGTTAAGTTTGTCTTTTCCAGAATCGAAGCAGTAAACTACAAATCGTTC
03/TCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCAAAGTTGTAAAGCTACAAATCGTTC
04/TCTTTAACCTCCTTGTTAAGCTTGTCTCTTGCAGAATCGAAGCTGTAAACTACAAATGCTTG

01/--CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG
02/TTCAAATGGAGCCCCAGATGCAGTCCATGAGTAAAATCTACCACGGACCCCTGGACCGGCCTG
03/TTCAAATGGAACCCCAGATGAAGTCCATGACTAAGATCTACCGTGGACCCCTGGACCGGCCTA
04/TTAAAATAGAGCCCCAGATGCAGTCCATGGCTAAGATCTACCACGGACCCCTGGACCGGCCTG

01/CTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCAC
02/CTAGCCCATGCTCTGATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCAC
03/CTAGCCCATGCTCCAATTGTAATGATATCGAACGCACCCCTCCCGAGGAAATCTCAACTGCAC
04/CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACGGCTTCCGAGGAAATCTCAACTGCAC

01/AACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCC
02/AACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCC
03/AACCCCTACTATGCCCCAATTCGCGAGGAAGCAGTTAGACTGGTCGTCAGCCAACCTCCCC

04/GACCCCTACTACACCCAATTTAGCGGGAAGCAATTAGAGCAGCCTATGGCCACCTCCCC

```

FIGURE 8.3

10/64

CTTCCCCAACTAATAAGGACCCCCCTTTCAACCCAAACAGTCCAAAAGGACATAGACAAAGGA 3
 CTTCCCCAACTAATAAGGACCCCCCTTTCAACCCAAACAGTCCAAAAGGACATAGACAAAGGA 4
 CTTCCCCAACTAATAAGGACCCCCC-TTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGG 5
 CTTCTCCAATAATAAGGACCCCCC-TTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGG 6
 CTTCCCCAAATAATAAGAACCCCCC-TTCAACCCAAACGGTCCAAAAGGAGATAGACAAAAGG 7

GTAAACAATGAACCAAAGAGTGCCAATATTCCCTGGTTATGCACCCTCCAAGCGGTGGGAG-- 3
 GTAAACAATGAACCAAAGAGTGCCAATATTCCCTGGTTATGCACCCTCCAAGCGGTGGGAG-- 4
 GTAAACAGTGAACCAAAGAGTGCCAATATTCCCAATTATGACCCCTCCAAGCAGTGGGAGGA 5
 GTAAACAATGAACCAAAGAGTGCCAATATTACACGATTATACTCGCTCCAAGCAGTGGGAG-- 6
 GTAAACAATAACCAAAGAATGCCAATATTCCCCGATTATGCCCCCTCCAAGCGGTGGGAG-- 7

A-AGAATTCGGCCCAGCCAGAGTGCATGTACCTTTTTCTCTCTCAC-ACTTGAAGCAAATTAAA 3
 A-AGAATTCGGCCCAGCCAGAGTGCATGTACCTTTTTCTCTCTCAC-ACTTGAAGCAAATTAAA 4
 AGAGAATTCGGCCCAGCCAGAGTGCATGTGCCTTTTTCTCTCCCAG-ACTTAAAGCAAATAAAA 5
 -GAGAATTTGGCCCAGCCAGCGTGCATGTACCTTTTTCTCTCTCAG-ATTTAAAGCAAATTAAA 6
 -GAGAATTCGGCCCAGCCAGAGTGCACGTACCTTTTTCTCTCTCTAGACTTTAAA----TTAAA 7

ATAGACNTAGGTNAATTNTCAGATAGCCCTGATGGYTATATTGATGTTTTACAAGGATTAGGA 3
 ATAGACXTAGGTXAATTXTCAGATAGCCCTGATGGXTATATTGATGTTTTACAAGGATTAGGA 4
 ACAGACTTAGGTAAATTCTCAGATAACCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGA 5
 ATAGACCTAGGTAAATTCTCAGATAACCCTGATGGCTATATTGATGTTTTACAAGGGTTAGGA 6
 ATAGACCTAGGTAAATTCTCAGATAACCCTAATGGCTATATTGATGTTTTACAAGGTTTAGGA 7

TTCTGAGTTCTTGCACTAACCTCAAAT 1
 CAATCCTTTGATCTGACATGGAGAGATATAATATTACTGCTAAATCAGACGCTAACCTCAAAT 3
 CAATCCTTTGATCTGACATGGAGAGATATAATATTACTGCTAAATCAGACGCTAACCTCAAAT 4
 CAATTCTTTGATCTGACATGGAGAGATATAATGTCACTGCTAAATCAGACACTAACCCCAAAT 5
 CAATCCTTTGATCTGACATGGAGAGATATAATGTTACTGCTAAATCAGACACTAACCCCAAAT 6
 CAATCCTTTGATCTGATATGGAGAGATATAATGTTACTGCTAAATCAGACACTAACCCCAAAT 7

GAGAGAAGTGCCGCATAACTGCAACCCAAGAGTTTGGCGATCCCTGGTATCTCAGTCAGGTC 1
 GAGAGAAGTGCTGCCATAACTGGAGCCCAGAGTTTGGCAATCTCTGGTATCTCAGTCAGGTC 3
 GAGAGAAGTGCTGCCATAACTGGAGCCCAGAGTTTGGCAATCTCTGGTATCTCAGTCAGGTC 4
 GAGAGAAGTGCCACCATAACTGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTC 5
 GAAAAAAGTGCTGCCATAACAGCAGCCTGAGAGTTTGGCGAATCTGGTATCTCAGTCAGGTC 6
 GACAGAAGTGTCGCCGTAACCTGGAGCCCAGAGTTTGGCAATCTCTGGTATCTCAGTCAGGTC 7

AATGACAGGATGACAACAGAGGAAAGATAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGT 1
 AATGATAGGATGACAACGGAGGAAAGAGAACGATTCCCCACAGGGCAGCAGGCAGTTCCCAGT 3
 AATGATAGGATGACAACGGAGGAAAGAGAACGATTCCCCACAGGGCAGCAGGCAGTTCCCAGT 4
 AATGATAGGATGACAACAGAGGAAAGAGAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGT 5
 AATGATAGGATGACAACAGATGAAAGAGAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGT 6
 AATGATAGGATGACAACAGAGGAAAGAGAACGATTCCCCACAGGCCAGCAGGCAGTTCCCAGT 7

GTAGACCCTCATTAGGACACAGAATCAGAACATGGAGATTGGTGCCGCAGACATTTGCTAACT 1
 AACT 2
 GTAGCTCCTCATTGGGACACAGAATCAGAACATGGAGATTGGTGCCGCAGACATTTACTAACT 3
 GTAGCTCCTCATTGGGACACAGAATCAGAACATGGAGATTGGTGCCGCAGACATTT 4
 CTAGACCCTCATTGGGACACAGAATCAGAACATGGAGATTGGTGCTGCAGACATTTGCTAACT 5
 GTAGACCCTCATTAGGACACAGAATCAGAACATGGAGATTGGTGCCGCAGACATTTGCTAACT 6
 GTAGACCCTCACTGGGACACAGAATCAGAACATGGAGATTGGTGCCGCAGACATTTGCTAACT 7

FIGURE 9.1

11/64

TGCGTGCTAGAAAGGACTAAGGAAAACCTAGGAAGA----	1
TGCGTGCTAGAAAGGACTAAGGAAAACCTAGGAAGA---CTATGAATTATTCAATGATGTCCACT	2
TGCGTGCTAGAAAGGACTAAGGAAAACCTAGGAAGA---CTATGAATTATTCAATGATGTCCACT	3
TGTGTGCTAGAAAGGACTAAGGAAAACCTAGGAAGAAGTCTATGAATTACTCAATGATGTCCACA	5
TGCGTGCTAGAAAGGACTAAGGAAAACCTAGGAAGAAGCCCATGAATTATTCAATGATGTCCCCT	6
TGCGTGCTAGAAAGGACTAAGGAAAACCTAGAAAGAAGCCTGTGAGTTATTCAATGATGTCCACT	7
ATAACACAGGGGAAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAG	1
ATAACACAGGGGAAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAG	2
ATAACACAGGGGAAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAG	3
ATAACACAGGG--AAGGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAG	5
ATAACACAGGG--AAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAAGGATTGAG	6
ATAACACAGGG--AAAGGAAGAAAATCCTACCGCCTTTCTGGAGTGACTAACGGAGGCATTGAG	7
GAAGCATACC---AGGCAAGTGGACATTGGAGGCTCTGGAAAAGGGAAAAGTTGGGAAAAGTA	1
GAAGCATACC---AGGCAAGTGGACATTGGAGGCTCTGGAAAAGGGAAAAGTTGGGCAAATTG	2
GAAGCATACC---AGGCAAGTGGACATTGGAGGCTCTGGAAAAGGGAAAAGTTGGGCAAATTG	3
GAAGCGTGCC232AGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAATTG	5
GAAGCATACC238AGGCAAATGGACTTTGGAGGCTCCAGAAAAGGGAAAAGCTGAGCAAATTG	6
GAAGCATACC233AGGCAAGCGGACTTTGGAGGCACTGGAAAAGGGAAAAGCTAGGCAAATCA	7
TATGTCTAATAGGGCTTGCTTCCAGTGTGGTCTACAAGGACACTTTAAAAAAGATTGTCC-AA	1
AATGCCTAATAGGGCTTGCTTCCAGTGCAGTCTACAAGGACGCTTTAGAAAAGATTGTCC-AA	2
AATGCCTAA	3
AATGCCTAATAGGGCTTGCTTCCAGTGCAGTCTACAAGGACACTTTAAAAAAGATTGTCC-AA	5
AATGCCTAACAGGGCTTGCTTCTAGTGTGGTCTACAAGGACACTTTAAAAAAGATTGTCC-AA	6
AATGCCTAATAGGGTTTGCTTCCAGTGCAGTCTACAAGGACACTTTAAAAAAGATTGTCCAAA	7
-TAGAAATAAGCCACCACCTCGTCCATGCCCCCTTATGTCAAGGGAATCACTGGAAGGCCCACT	1
GTAGAAATAAGCCGCCCC-TCGTCCATGCCCCCTTATGTCAAGGGAATCACTGGAAGGCCCACT	2
GTAGAAGTAAGCCGCCCCCTCGTCCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACT	5
GTAGAAACAAGCTGCCCCCTTGTCCATGCCCCCTTATGTCAAGGGAATCACTGGAAGGCCCACT	6
-TAGAAATAAGCCGCCCCCTCGTCCATGCACCTCGTGTCAAGGGAATCACTGTAAGGCCCACT	7
GCCCCAGGGGATGAAGGTCTCTGAGTCAGAAGCCACTAACCAGATGA	1
GCCCCAGGGGACGAAGGTCTCTGAGTCAGAAGCCACTAACCCTGATGA	2
GCCCCAGGGGACAAAGGTCTCTGAGTCAGAAGCCACTAACCAGATGA	5
GCCCCAGGAGATGAAGGTCTCTGAGTCAGAAGCCACTAACCAGATAA	6
GCCCCAGGGGACGTAGGTCTCTGAGTCAGAAGCCACTAACCAGATGA	7

FIGURE 9.2

12/64

RTPLSTQTVQKDIDKGVNNEPKSANIPWLCTLQAVGEEFGPARVHVPFSLSHLKQIKIDG SDSPDG
- = == ===== = ===== = ===== ===== =====
KDPPSTQMVQKEIDKRVNSEPKSANIPQLPLQAVGGREFGPARVHVPFSLPDLKQIKTDLGKFSDNPDG

YIDVLQGLGQSFDLTWRDIILLNQTLSNERSAAITGAREFGNLWYLSQVNDRTTEERERFPTGQQ
===== ===== ===== =====
YIDVLQGLGQFFDLTWRDIMSLNQTLPNERSATITAA~~X~~EFGDLWYLSQVNDRTTEEREX~~X~~FPTGQQ

AVPSVAPHWDTESEHGDWCRRHLLTCVLEGLRKTRK TMNYSMMSTITQ GK
===== -----
AVPSLDPHWDTESEHGDWCRRHLLTCVLEGLRKTRK KSMNYSMMSTITQ GR

FIGURE 10

[illegible]

FIGURE 11

14/64

```

ACCTTGCAAGATCAACTTA- ACTCCCTAGCAGCAGT- AGTCCTTCAAAATCGAAGAGCTTTAGACTTGCT
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
ACTTTACAATCCCAAATAAGACTCTTTGGCAGCAGTGACTC- TCCAAAACCGCTGAGGCCTAGATCTCCT
AACCGCTGAAAGAGGGGGAACCTGTTTATTTTATAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGA
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CACTGCTGAAAAAGGAGGACTCTGCACCTTCTTAGGGGAAGAGTGTTGTTTTTACACTAACCAGTCAGGG
ATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATA- -CAACGTAGAGCAGAGGAGCTTCGAAACAC
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
ATAG-CA-TGAGAT-GCCACCCAGCGTTTACAG- GAAAAGGCTTCTGAAATCAGACGCCTTTC- AAATTC
TGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCTTCTTAGGACCTCTAGCAGCTATA
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
TTATACCAA--CCTCTGGAGT--TGGGCAACATGGCTTCTCCCTTTCTAGGTCCCGTGGCAGCCATC
ATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAG
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
TTGCTGTTACTCGCCTTTGGGCCCCGTATTTTAAACCTTCTTGTCAAATTTGTTTGGTCTAGAATCGAGG
C--T--G-TAAA-A---CT-ACAAATGGAGCCCAAGATGCAGTCCAAG- ACTAAGATCTACCGCAGAC
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CCATCAAGCTACAGATGGTCTTACAAATCGAACCCCCAAATG- AGTTCAACTAACAACCTTCTACCGAGGAC
CCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAG- GCACCCCTCCTGA- GGAAATC
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CCCTGGACTGACCAGCTGGC--ACT-TCCCCCTG----GCC-T-AGAGAGTTCCCCCTC- TGAAGGACA-C
T-CAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCC
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
TACAACTGCAAAGCCCCTTCTTCGCCCCTATCCAGCAGGAAGTAGCTAGAGCAGTCATCGGCCAAATTCC
CCAACAGCACTTAGGTTTTCTGTTGAGATGGGGG
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
C-AACAGCAGTTGGGGTGTCTGTTGAT-TGAGGG

```

FIGURE 12

15/64

```

agttgcaattccttgccctcaactctgagagaaacccagccacatctccagcaaacaaga
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
agttgcaattccttgccctccactgtgagacaaacccagacacatctccagcacacaaga 2299

acttcaaaacacctgaactgcagcagccaggcggttcctccaggaccacctccccaggat
||||| ||| ||| ||| ||| ||||| ||||| ||||| ||| ||||| |||||
acttcgaaatgcctcaacctcaggtgccaggggttcctccagaaccttctccccaggag 2359

cttgcttcaagtgccggaatctgaccattggggccaaggaatgcctgcagcccaggattc
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
cttgctacaagtgcagaaatctggccactggggccaaggaatgccacagaccaggattc 2419

ctcctaagccacgtcccatttgtgcaggacccactggaaatcggaactgtccaactcacc
||||||| ||||| || || ||||| ||||| ||||| ||||| ||||| |||||
ctcctaagctgtatcccatctctgtgggacccactaaaaatcagactgttcaactcacc 2479

cggcagccaatcccagagcccctggaactctggcccaaggctctctgactgactccttcc
||||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
tggcagccacttccagagcccctggaactctagcccaaggctctctgactgaccttct 2539

cagatcttctcggcttagcagctgaagactgacactgcccgatcacttcagaagtccct
||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
gagatcttcttggcttagcagctgaagactgacactgccagatcgcttcggaagcctaca 2599

ggaccatcacggatactgagcttcaggtaactctcacagtggaggctaagtccatccct
||||||| ||| || ||||| ||||| ||||| ||||| ||||| |||||
ggaccatcacagat-----gctccaggtaactctcacagtagagggttaagtctgtccct 2654

gtttaatcgatacaggggctacccactccacatcaccttcttttcaagggcctgtttccc
||||| ||| | ||||| ||||| ||||| ||||| ||||| ||||| |||||
tcttaatcaatatggaggctacccactgcacattaccttcttttcaagggcctgtttcct 2714

tttccccataactgttgtgggtattgacggccaagcttcaaaaccccttaaaactcccc
|| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
ttgctccataactgttgtgggtattgacggccaggcttctaaacctcttaaaactcccc 2774

cactctggtgccaaacttggaacattcttttatgcactctttttcagttatcctcacct
||||| || ||||| ||||| | ||||| ||||| ||||| ||||| ||||| |||||
aactctagtagcaacttagacaatactcttttaagcactcctttttagttatccccactt 2834

gcccagttcccttattaggccgagacattttaaccaaattatctgcttccccgactattc
||||||| ||||| ||||| || || || ||||| ||||| ||||| ||||| |||||
gcccagttcccttatgaggccgagacacttcaactaaattatctgcttccctgactattc 2894

ctgggctacagccacatctccttgccgcccttcttcccaacccaaagcctccttcatatc
|||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
ctggactacagctacatctcattgctgccttcttcccaatccaaagcctcctttgcatc 2954

ttcctctcatatccccccaccttaaccacaaagtatgggacacctctactccctccctgg
||| | ||||| ||||| ||||| ||||| || ||||| ||||| ||||| |||||
ttcttgt---atcccccaaccttaaccacaaagtataagatacctctattccctccttgg 3011

```

FIGURE 13.1

caaccgatcacacgcccattactatcccattaaaacctaatacacccttaccctgctcaat
||| |||| | ||| |||| || |||||||||||||||| |||||| ||||||
tgaccaatcatgcaccccttaccatctcattaaaacctaatacactcttacccggtcaat 3071

gccagtatcccataaccacaacaggctttaaaagggttgaagcctgttatcacttgccctgc
|||| |||||| |||||| || |||||| |||| || |||||| |||||| ||||||
gccaaagatcccaccccacagcatgctttaaaaggattaaaacctgttatcactcgccctgc 3131

tacagcacgggcttctaaaacctataaaactctccatacaattcccccatTTTtacctgtct
|| |||| | || ||| |||||| |||||| |||||| |||||| |||||| ||||||
tagagcatggcctTTTtaaagcctataaaactctccttacaattcccccatTTTtacctgtcc 3191

aaaaaccagataagtcttacagggttagttcagaatctgcaccttatcaaccaaattgttt
| |||||| || |||||| || || |||| || || |||| || || |||| || ||||
tagaaccagacaagccttacagggtt----caggatctgtgtcttatcaatgaaattgttt 3247

tgctatccaccctgtagcacccaactcgtacactcttttgcctcaatgccttccccca
| |||||| |||||| || | ||| || || |||| | |||||| || || || || || ||
tccctatccaccctgtggtgctgaaccatatactctcctatcctcaatacctccctcta 3307

caactcactattccgttcttgatcttaaagatgcttttttactattccccctgcacccct
|||| | |||||| |||||| |||| || |||||| || |||||| || || || || || ||
caaccattattctgttctagatctcaaacatgctttctttactatccctttacaccctt 3367

catcccagcctctctttgcttttacctggactgacccctgacacccatcagtcccagcagc
|| |||||| |||||| || ||| |||||| |||||| |||||| |||||| ||||||
caaccagcctctcttcgttttaccctggactgacccctgacacccatcagtcccagcagc 3427

ttacctgggctgtactgccgcaaggcttcagggacagccctcattacttcagccaagctc
|||||||| |||| || |||||| || || |||||| || || |||||| || || ||||||
ttacctgggctgtaatgctgcaaggcttcaggggagcccttattatttcagccaagctc 3487

tttctcatgatttactttctttccacctctctgcttctcacctattcaatatattgatg
|||||| |||||| |||||| || || |||||| |||||| |||||| |||||| ||||||
tttctcatgatttactttctttccaccctccacttctcacctattcaatatattggtg 3547

accttctactttgtagccctccttttaaactcttctcaacaagacacccctcctgctccttc
| |||| |||||| |||||| |||||| |||||| |||||| |||||| || || ||||||
atgttcttctttgtagccctcctttgaatcttctcaacaagacacacttctgctccttc 3607

aacatttggttctcaaaggatatcggtatccccctcaaagctcaaatttcttctccat
| |||| |||||| |||||| || || |||||| |||||| |||||| || ||||||
agcatttatttctcaaaggatatc-----ccctcaaagctcaaattgcttctccat 3660

ctgttacatacctcgggcataattcttcatgaaaaacacatgtgctctccctgccaatgcg
| |||| |||||| |||||| || || |||||| |||||| |||||| || || ||||||
cgtttacctaccttgggcataattcttcataaaaaacacacgtgccctccctgctgatagtg 3720

tctccaactgatctctcaaatacccaacctcttctacaaaaacaacaactcctttccctcct
||| |||||| |||||| |||||| || || |||||| |||||| |||||| || ||||
tctg--actgatctctcaaaccccaacccttctacaaaaacaacaactcctttccatcct 3778

aggcatggttgatacttttgcctttggatacctggttttgccatcctaacaaaaatcatt
|||| |||||| |||||| || || |||||| |||||| |||||| || || ||||||
aggcatggttgatactttcgtgttaggatacctggttttgccatcctaacaaaaaccatt 3833

FIGURE 13.2

17/64

atataaaactcacaaaaggaaacctagctgaccccatagattctaaatcctttcccccactc
 |||||
 atataaaactcacaaaaggaaacctagttgaccccatagatcctaaatcgtttcccccactc 3898

ctctttccattccttgaagacagcttttagagactgctcccacactagctctccctgtctc
 |||||
 ctctttccattccttgaagacagcttttagagactgtctccactctagctctccctgactc 3958

atcccaaccctttttcattacacacagccgaagtgcagggctgtgcagtcggaattcttac
 |||||
 atcccaacactttttcattacacacagctgaagtgcagggctgtgcagtcagaattcttac 4018

acaaggaccgggacccatgcctgtagcctttttgtccaaacaacttgaccttactgtttt
 |||||
 acaaggaccgggatcgcatcctgtagcctttttgtccaaacaacttgaccttactgtttt 4078

aggctcgccatcatgtctccatgcggtagcttccgctgccctaataacttttagaggccct
 |||||
 aggctggccatcatgtctccatgcagcgtctgtgccaccctaataacttttagaggccct 4138

caaaatcacaaactatgctcaactcactctctacagctctcacaacttccaaaatctatt
 |||||
 caaaatcacaaactatgctcaactcattctctacagctctcataatttccaaaatctatt 4198

ttctttctcacacctgacgcataatactttctgtccccggctccttcagctgtattcact
 |||||
 ttcttcctcacacctgacacataatactttctgtccccggctccttcagatataactcact 4258

ctttggtgagtctcccacaattaccattcttctctggcccagacttcaatctggcctccca
 |
 c--catttattctcccacaattaccattattctctggcctggacttcaatccggcctccca 4316

cattattctggataccacacctgaccctgatgattgtatgtctctgatctacctgacatt
 |||||
 cattattctggataccataacctgaccctcatgactgcatctctctgatccacctgacgtt 4376

caccctatttccccatatttcttcttttctgttctcatgttgatcacatttggtttac
 |||||
 caccctatttccccacatttcttctgtccctgtttctcaccctgatcacacttggtttat 4436

tgacggcagttccaccaggcctgatcgccactcaccagcaaaggcaggctatgctat
 |||||
 tgatggcagttccaccaggcctaatacgccactcaccagcaaaggcaggatagctat 4493

gaactgattgccttaactcgggccttctactcttgcaaagggactacagtcgaatatttat
 |||||
 gaactagttgccttaattcaagccctcactcttgcaaagggactacgtgtcaatatctat 4553

actgactctaaatatgccttccatatcttgccaccacatgctgttatatgggctgaaaga
 |||||
 actgattctaaatatgccttccatatcttgccaccacatgcggctcatatgggctgaaaga 4613

ggtttcctcactacgcaagggtcctccatcattaatgcctctttaataaaaaactcttctc
 |||||
 ggtttcctcactacacaagtgtcctccatcattaatgcctctttaagaaaa-ctctgctc 4672

FIGURE 13.3

aaggctgctttacttccaaaggaagctggagtcacacactgcaagggccaccaaaaggcg
|||||
aaggctgctttacttccaaaggaagctggggtcattcactgcaaggggcatcaaaagact 4732

tcagatcccattactctaggaatgcttatgctgataaggtagctaaagaagcacctagc
|||||
tcagatcccattgctctaggaatgcttatgctgataaggtggctagacaagcagctagc 4792

gttccaacttctgtccctcatggccagtttttctccttcccatcagtcattcccacctac
|||||
tctccaacttttgtccctcatggccagtttttctccttcacatccgtcactcccacctac 4852

tccccattgaaacttccgcctatcaatctcttctcacacaaggcaaatggttcttagac
|||
tccacagctgaaacttccacctatcaagctcttcccccgcaaggtaaatggttcttagac 4912

caaggaaaatatctccttcagcctcacaggccattctattctgtcatcatttcataac
|||||
caaggaaaatatctccttcagcctcacaggccattctattctgtcgtcatttcataac 4972

ctcttccatgtaggttacaagccactagtcacctcttagaacctctcatttcctt-cca
||
cttttccatgtaggttacaagccactagcctgtctcttaggacctctcatttcctttcca 5032

tcgtggaaacatatcctcaaggaaatcacttctcagtgttccatctgctattctactacc
||
tcatggaaatctatcctcaaggagatcacttctcagtgttccatctgctattctgctacc 5092

cctcagggattgttcaggccccctccctccctacacatcaagctcggggatttgccct
|||||
cctcagggattgttcaggcctcctcccttctcctacacataaagctcggggatttgccct 5152

gccaggactggcaaattgactttactcacatgcctgagtcaggaaactaaaatacctc
|||||
gccaggactggcaaattgactttactcacatgcctcggtcagaaaactaaaatatctc 5212

ttggtctgggtagacactgtcactggatgggtagaggcctttcccacagggtctgagaag
||
ttagtctgggtagacactttcactgggtgggtagaggcctttcccatagagtctgagaag 5272

gccactgcagtcatttcttcccttctgtcagacataattccttgggttgcccttcccacc
|||||
gccaccgcggtcatttcttcccttctgtcagacataattccttgggttgcccttcccttc 5332

tctatacagtcctaataacgggagcagcctttattagtc aaatcacctgagcagtttttcag
|||||
tctatacagtcctgataaacggaccagcctttactagttaaatacccaagcagtttctcag 5392

gctcttgggtattcagtggaaaccttcgtaccccttactgtcctcaatcttcaggaaaggta
|||||
gctcttgggtattcagtggaaaccttcataatcccttaacatcctcaatcttcaggaaaggta 5452

gaatggactaatggtccttttaaaacacacccccaccaaactcagcctccaacttaaaaag
||
aaaccgactaatggtccttttaaaagacacacctcaccaagctcagcctccaacttaaaaag 5512

FIGURE 13.4

19/64

TGCCTTTATTTCCGTAGGCTGGTCATATGGCGCTAGCACTCACATAAAGCTACCGAGGAG
 AGCGAATGAAACCAAAATCACTTTACCTTCACAGCACGAGGCCGTCGTCCCTCTCGATAT
 TTGGCCCGTGTGTGTCGATACCGCCCTCTGGACGTGGTGATCAAATAAACTCCCTAGCTCC
 CCGCCGCTCGACGCCATCTTGCCTACTTTGATCCTCGCAGGGAGGACAACATCCGCCCTA
 CTGAGCTCCCTTTTATCCAATAAGAGAGCGGGATGAGTTAAGGAGTGCCAGGATTGGCTG
 GAGAATCGACAGCGTCGGCCATCGTTTCCTGCGTGCGAAGATTTGATGAACGAGGTGCCG
 CCCCCGAGCGGCTCGGCGGAGAGGCGCGGTGGGTGACAGAAGCTTTCTTGTCCCACCCAC
 TACAGGCTTACGGCAGGATGCGCAGCGGGGAGAGGGGCGGGGCCGAGGGGGCGGGGCC
 GATCGATCTCCTCCGGCTCCGACGTCTCGGCCTGCCGGGTCCCGGGTCTTTGCGGCGC
TAGGGTGGGCGAACCAGAGCGACGCTCCGGGACGATGTGGGGCAGCGATCGCCTGGCGG
GTGCTGGGGGAGGCGGGGCGGCAGTGACTGTGGCCTTACCAACGCTCGCGACTGCTTCC
TCCACCTGCCGCGGCGTCTCGTGGCCAGCTGCATCTGCTGCAGGTAACCTGCCGGCCCC
 GAGCCACCTGATCTTCAGCCTGGGGTCGGACGAGGCCGAAGCCTCTCAGGGACGCGGCGG
 GACACCGGCTGCCACCCGGGCGCCGGAAGCGCGCAGAGATCAGGGTCCCTCGACGCCA
 GGCCCTTCTGGGTAGTCTCTGGATCCCAAGTCCAGTGCAGCCCTGGGCTCGTCTTAT
 CCCAGGTCTTTTCACTTGGTGAAACTGAACCTAGAAACGTCTTAATATTCTACCACTGTT
 TTTATAAATATTCTTATTCCAGGCTGGAAAAGCTCCTGAGAAGTGGTTTTGTTTTTATTA
 TTTTAAAGGTGTTTTCTTGGCAGCCATTTCCAGTTAACCTGCGCTGCTGCCGTCCGGG
 CCGCGAGAGCGGGACGCAGAGTTGTTGGCGGAGCCCCTGTCGGTTCCCGGGGACTAAGCA
 CCGCGTCCCATGAGCGGGAAAGGTTAATAACAATGATGGTTCTGCCCTGCGTCGCTGACGC
 GGAACACAGCTGTAGTGTGTTAGGAACACATAACGTAGTTAAGATCACTTGAAGCTCTGC
 GATCAGTCGCCCTTCTGGACGTTGTGGTTAGGATGTTTCACAGTTCTAACCCTGGTGGA
 GATACAGCGTCCATATTTTATATAATTAATAAGAGGCACATGGTCTCACGAGTTTGAGT
 GTACTTATGGGGGCAAAAGGACGGCGTATTTGAAATCCTCATAAATCCTGGATGCATGGT
 ACCCACCAGTGGCTAATCTATGCAATGAATAGAGTTTGCAATAATTTCAAGCATCCCTTC
 TTTCCACTTGAGTTACTTCCCCATACCTAGGGGAAGATATTTTTTGGTCCACTGAAAACAT
 GAGTTCAGCAGAATCCTCCTATCATCGTCGTTATTATTTTTTACCACTAAGTAGACAATC
 TTTTGGTTTTTGGATGGGCTTTATGGCTAGAGACAAATCAGTCACTGTCACCAAGTTCAG
 GTAGAAGTTGGTTCACTGCTCTGTGTCAGCTTCGATGGGATTTTTCAACATGTTTTCAAATC
 TGCACCTAATAGTAGGAATGCTTTCTTACAGTAACTCTAATTTGATCCTAAGATGTAGTT
 GTTACCTTACATTCATCACTGTTTAAAGAATTTAGTGGTCTTGATCTTTGTTTTAAATTTT
 GAGCCTTCGGGAAGTACTTATAAGAATTAATTCATGCATATCTTTTTGAAATGTAAATGT
 CTTTACGCCCTGGAACAATTGCTGTTTCTGTTTCTAGCCCATATTAGCAGAATAGGTCAACT
 TTACTTTCTAATTAATAAGTAAATAGTTTATTACTTTATAGATTCCATAAATCTATACA
 TTTATTCCTCGATGAATTATATAAATTTATAGAATTTATGTTTTATAGAAAATTTGGAAA
 GCATGGAAAATTATTAACAAGAAAATAAGTTACCCATAATCCAGAACTTAGAGGTGACT
 AATGTTGACAGTTTGGATCAAATCTTCCAGTTTTGTTTCTAATCTTTATTTTAAACATAA
 ATGAGGTCCTGTATACACACGTACAGTTTTGTGTCCTGGTGTTTTTATTAAATGTTATTA
 TGAGTGTTTTTATTTTGTAAAGGTCATCATTTTAAAGTTGTTAATTAGTATTCTAGCACA
 AATTTGCCATAATTTATTTAATTGTTTACTATGATTGACCATTTAGATTGTACTTAATTT
 TTAGGCATTAGAAGTGATAAACTATATTTTAAATCAGACGTTGAAAATAACACATCTTTGT
 TTAGAAAACATCATTTTATTTCTGGTTGTCTAGGATAGATTCCCAGAATCTTTGGGTTAG
 AGGCCATAGATAATTATGAAAGCAGAAAGATTCAAGTTGGGAGTTAATACTTGAATTA
 CTTTATTTGGGGTGAAGCATTGAGTGCATAATACAGATCATGCAGTAATGGGAAGAAGGG
 TTGGAACAATGGTTTTCTGGCCTATGTCAGACTTACCTTGAAGCTTTTAAAGAATACAGAT
 GTTCTGATCAACCCTCAGACCTATTAAATCAGACCTAAAATCTTAGGGAATAGGCTTTAG
 GCATCTCTAATTTTAAAAAATTTATTTCAGGCTACTTGGATGCACAAAAGAGTTGAGACCT
 ACTGTCCTAGAATCATAGAATTTTAAATGACGATAGAGACCTTAAAGCATCTAGGTCGTTTC
 TGTACTTTTACATGTAAGGAACTGGCATTCTAGGCCAGTACCATTGCCATGCAGCTAA
 TTTGCCCTCTTGTCTATAGCTCACTCTGCATCACCCAACTACCGTTCTCACTGTTTCTT
 CTATAACCAATCTCCTTCCCCTTCTGTTCTTACTCATGCCATTCTTCCCTCAGTCAT
 TTTTCTTCTTCCATACAAATCCATGTCTTTAAAAAGGAATAATCCTACCTCCTCCACA

FIGURE 14.1

20/64

TAGCTTTCCAATTCTCTGTTGCCACATTTGTCTCCCTTTCAATACTTCTCTGTTGTGTT
 ATGTGACACATCACATTTGATATACTCTGTACTGTGTTTCAAGTATTGTATTCTCTTGTT
 TACTCAAGTCATTATTTTCAAGTACTGACTACCCAGTAGATGCTTTAAGTCAGGATTTCTCA
 ACCTTGGCACTGTTGACATTTTGTAGCTGGATAATTTTTTGTGTTTGGGGGCTCTCCTGTAC
 ATTTTAAGATGTTTAAACAGCACCCCTTGGCCTCTATCCAGTAGACGCTGTACTGCCTCCC
 CCTATCTGTGACAACCAAAAAGGTCTTCAGACATTGTCAGATGTCTACTGAAGGACAAAA
 TCACCTCTGGTTGAGAACCACCGCTTCAACTAAGTTATCTTCTCTGTACTCAGAACTTGA
 TGTGATTGCAGCAGGGGGAGAGGATTCATATACACAGTGAATGCAAACGAACCTAAATCA
 CCATTTCGGATATGGCCACACAATTTTCATTTCCCTTGTGTTAGCAAGAGATACCTTAGGC
 TTTGGACCTGATTATTCTTAAGGCATTCTGATGTATGGTTTTACCTGCAGATTTCTCTGGT
 AATACTGATACCTCAGTTTGGGTCAAAGAAGGTCAATTAATTGATTGATTTGATTTGACT
 CCTGGAAAAGACGCTCCTTTCTAGCTGTCTCTTTCTCTTTTACCTGAATAGCCAGGGC
 TCTGTGGTTCAAGTGAAGTATTTTGTACATAAAAAATTAACCTTAGAACATTGGTCTGCAGAG
 TTTGCTCAATATAACTGAGCACATATTGTGGCTTTATGGAGCTGGTTACTACTTTTGTAC
 CAAATAAATAATTAGAAGTATTTTCTCCTCAATAAGGTTTCAATTTTCTTTTTCAGT
 GAGCTGGTAGAGTTTCTTTTGTATTTTCAAGGCATCTTTCATATTTCCATCTCTTAA
 GTTTCTTTCATATGAAGTAGAATTTATCTGGATTATGTATTGCTGACTCTGATGAAAACCC
 ATAGAAGCATCTGGGGCTTGTATCACCTTCTTGTGTAATAGCTCACACGGTTACAGCT
 GATATGGTACTTAAGACTTTTGTATCCAAATCTAGGCAAAATACACTCAGTTGAAAAGAA
 TTTGTGAGCCAGAACAGTTGGACTGTTCTGTGAAAATTGTGAGAAAAATTACACAATAA
 GTGATACATGATGATGGCTTTCTTAAATATAAAATTGTAATAACATGGTTAATTTCCAGT
 ACGTTATATTGTCCCAGAAGTGGCTCCAACATTGTTTGAATTTGTCTCATTTAAAGAAA
 CATAAGCTGGCTATGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCAGG
 CAGATCACCTGAGGTGAGGAGTTGAGACCAGCCTGGCCAACATGGTAAAACCCCATCTC
 TACTAAAAATACAAAAATTAGCCGGGCATTTGGTGGGGGCTGTAATCCCAGCTACTTGG
 GAGGCTGAGGCAGGAGAATTGCTTGAATCTGGGAGGTGGAGGTTGCAGTGAGCCGAGATT
 GTGCCACTGCCCTCCAGCCTGGGTGACAGAGTGAGTCTCCGTCTCAAGAAAAAAAAAAAA
 AAAAGCAAGAAACATAAAGACTGGGCATGTTGGCTCATGCCTGTAATCCCAGCACTTTGA
 GAGACTGAGGTGGGAAGATCACTTGAGCCCAGGAGGTTAAGGCTGCAGTGAGCCGTGATT
 TTGCCACTGTACTCGAGCCTGGGCAACACAGTGAGATCCTGTCTCAGGAAAAAAAAAAATT
 GCATGTAAATGAATGAATTTGATATTTAATATTTTAAATTATGAAAACCTGTTCTGTAGAG
 ATGTAGATCTTGCCATGTTGCCAGGCTGGCTTTGAACTTCTGGGCTCAAACAATCCTCC
 TGTCTCAGTCTCCCAAAGTATAAAGATTACACATGTGAGCCACTGCACCTGGCCTAATAT
 TTTTAACTTAATGAATTTATTTTGTATATAAATAAATTAAATAACACTGAAGCTTCTGTATA
 TAATAAGTCTTTTGTGTGTGTGACGGGTCTCACTCTGTTGCCAGACTGGAGTGTAAAT
 GGCATATCATGGCTCACTGTAGCCTCAACCTCCCTGACTCAAGTGATCCTCCCACCTCG
 GCTTCTGAGTAGATGGGACCACAGGCGTATGCCACCACACCTGGCTGATTTTAAAAATT
 TATTATTGATACATATTAATAAAAATTATTTTATTTTAAAAATGATATATGTGGCTGGGC
 ATGGTGGCTCATGCCTGTAATCCCAGACGTTTGGGAGGCCGAGGTGGGAGGATCACTTGA
 GACCAGGAGCTTAAGACCAGCCTAAGCAACATAGTGAGATCCCATCTCTATAGAAAAAAA
 AAATGGCTAGGTGTGGTGGTGTATGCCTATATTTCCAGCTACTCAGGAGACTGAGGTGAG
 AGGATTGCTAGAGCCCAGGAGTTTCAAGTTACAGTGACCTATGATTGTGCCAGTGCACTC
 CAGCCTGGGCAACAGAGCAAAATCCTGTCTCAAAAAAAAAAAAAAGTTGAAAAATGCTTAT
 GATGCAATATAAGTAGTGAAAAGGATATTAAATTGTGCCTATATGAACACAACTATATG
 AAAAACTTGACATAGAGAAAAGGATTAACAAGAAATAGACCAAATTGTTTACATGGTTG
 TCTTGTGTTGTGGAGAGAATATCAGTAGTTTCAATTTGTTTCTTCCAAGTTTATATGTTTC
 CGAGGTCTCTATAATGAGTTTGTAAATTGTTTAAATCATAGAAAACCTTTTTTGGTCTCTG
 GCCACAACTTACATGTTTAAATGTAATTGCTTTTTTAAATGAGAATAAATGTTATATTTT
 GCTTTTTTAAACCTATATTCCCATAGTTATATGAGCCCTTACAATTATTAAGAGGCTGC
 ATAATATAACGTTTCTGGAAGGGTACAGAAGAAAACAGCAGTAATTACCTCTGAGAACAGA
 GACATGGCTTCACATTTTACCCTTTTGTACGTTTTGTGCTTTTGGCCACATGCATTTATTA
 TTCTTCCAATAAATAAGTAAATAAATATGGATTGTATACTCCATCTGGTTGGTGTTCAT
 AATTCTAAAAATTATATTGCTACATTTTTAAAGATGATATGTGTTTCTACTTATTAACGTA

FIGURE 14.2

21/64

TATGTTAAATAGTAAATTTATATCTTATTTAATAATTTCCCTATTGATAGACATTTAAG
 ACAGTCTCAAGTGTTCACCTATCATAGAAAATACTGCACAGATAGCTTTTGCTATAGTTTC
 TTTTTCTTTGAATCGTTAATTGGGAATAAATGCTCAAATAGTTATATGTGGCTCAACTG
 CTATTTAAGTTTTATTGACTGACTGCTGCCATTTTGAATTCTGAAGGGGTGATTAAATTT
 ATAATGCTGCCATAAGAATATAAGGGTATTGGCTTCATTAGCATCCACCAGCATTGGGTG
 TTGGAAATGATTATAGATTTTTTAAATGCTACAACAAATGTAGATAACAGAGAACATCTA
 TAGAACTCTTTTTGGACATGTGAATTGTAATAATAGTTTATTTTTCATGTGAATCCAGAAA
 AATGTATACGAAAACCTTTTTCTCTCATTTCTTATATGAATAGAAATCAAGCTATAGAA
GTGGTCTGGAGTCACCAGCCTGCATTCTTGAGCTGGGTGGAAGGCAGGCATTTTAGTGAT
GGGGGACAGGTAAGCACATGTGATGGCAATAACTTTCTCTAATATCACATAATATAGCA
 ATAGAAATAAAATTTAAAGTTTAGATTTTTTGTAAAGGAGGTGAGATGTCACCTAATTT
 GTATGCTATTATGTAAGTCTAGGATATTGAAGCTGACTATACTCTGTTTTTAGGTCA
 TTATCTTGTAGTTTACCATACTCCCTACTTGCTTCTTATTCTACTATTTAACTCATTTTC
 CACATCCCCTAATTTTTGGTTTTCATGAAATTATTTTTCTTCTGAATTACTAGGTTCTACT
 TACTATTATTAACTTTATTTCTGACATATTTTATAACCTTCCATGGTCTCACTTGATTA
 AAAATAAAAAAATTCAGCTGGGTGCGGTGGCTCACACCTATAATCCCAGCACTTTGGGAGG
 CCAAGGTGGGCGGATAATTTGAGGTGAGGAGTTGGAGACCAGCCTGCCCAACGTGGTGAA
 ACCCCCCCTCTCTACTAAAAATTCAAAAATTAGCTGGGCATGGTGGCAGGTGCCCTGTAAT
 CCCAGCTACTCAGGAGGCTGAGGCAGGAGAATTGCTTGAACCTGGGAGGTGGAGGTGCA
 GTGAGCTGAGATTGCACCTGCTGCACCTCAGCTGGGTGACAAGAGCGAAACAATGCTTTGA
 AAAAAATAAAAAATAAAAAATTTCTACAACACAGGGTTATTATTTTTCCATTTTTGTTTT
 CCCTTATGAGTTTAAATATGTTTAGATTATAAACCTGAAAGCTTGAATACCTATGCTATC
 TTTTGTTTTCTTATGTTTATCAAGTTATTCCCTTTAAACATTTTCTAAACTGTAAGAATAA
 TGTGAGGCTGGGCTCAATGGCTTATGCCCTGTAATCCCAGTGCTTTGGGAGGCCAAGGTGG
 GAGGACCCTTGAGGCCACGAGTTCAAGATTAGCCTGGCTAGGCAACATAGCAAGACCCT
 ATCTCTATAAAAAAATTTAAAAAATTAGCTGGGCATGGTAGCAAATGCTTGTAGTCCCAG
 CTACTCAGCAGACTGAGGTAGGAGGAATGCTTGAAGACCAGGAATTTGAGTGACCTATGAT
 TATGCACCTCCAGCCCGGGCAATAGCAAGACCCTATCTCTTAAAGAAGAAGATGTAGTAA
 TAATACATATTCTATTATAACTATTTTACCATTGAAAGTAAAAATGAGTTTTTACCTTTT
 CCCAGTCCCATCCTCAGAATGGGGATCTCAGTAGACCTTTAGGATTGGAAGAATGAGATC
 ATTCATATTTTCTGCAATTATTACCCACAAAATATTTTCAAGATACCTTTCCATGTATTAC
 AAACAATGTGCATTTAACATGTCTCTCTCTCTCTCTCTCTGTGTGCGTCTTCATGA
 TCCTCTGTTGCAGCCCTGCCAGTAAGACACTATCTCCTGAAGAATCACTGATAGGAACAG
 AAAGTGGACTGGCTAGGCCAGGAGTCCCTAGCTTCTTAGGGGGCAGGAGCTGCTTTGTGC
 TTTCTCAGAATCAGATATATATGTGGACTGAAACATTTAAAAACAGAATAGCCAAGGGTG
 CTATACGTTTTAAACCTTATATAGATGGGGCTACATTGCTCTCTATTACTAATTTCCCATG
 ACAATACACGAGAGTGCCATGTCTTTTAACTTGTTTTGAAGCACAGACTAATCTTGTTTA
 TGCATGTTTTTTGATGAGAATAGGCTACTCATGAGAAATCTGTAAACCTAACACTAGTCC
 CTTGCATACTCTAAATTTGTTGCTAGAATCTTAAATTTTAGCACCAGACGGACCTTAGAA
 ATCATTAACCTTTGGTGTCTTTGTTCTACAATAACAAGGAGATGGAATATTTTACCAGGATT
 GCTTAGCAGGTTACAGTTCTGCCCTCTGAGTACCCAGCACTTCCCTGTGGGCAACATCAA
 CTTCCCTGATTTTCAAGTCTTAATTAGTACTCTGAAGAATCCTACTTGTTTTTTAACCTCCA
 TTTGCTTTGAAGTGACTTTACCTGATTTTTTTTAGATCCCTTATTGCAGCAATGCCACTAA
 GAAACTGAGTCTCTAGCTTCTTGGTGGGCAGGAGCTGCTTTGTGCTTGCTCAGAATCATC
 CTTTTCAGTAAGGGAGATATTGAAGAGAAATCTACTGAGGAGTCTGGGGGTGAGGCACCTC
 AGGGAAATCCTGCTCCAGTCCACAAAAGCAGAGAGGAAGGGTTGGTTACCTAGAGTATTT
 AACATGCAGAGGCTTTGGATTTTACTCCTTTAATCCTTGGAAATGCCTATGGAAGGGGAA
 AGGAAGTAAGATGGTGAAGTCCAGCTTATAGACATACTAGTGTACATATATTTAACTAT
 AATAGGAGGGTATTATTAGTTTTACTTAACTTTCAACTGTGAAGGATTATACTTCTCAAT
 ATTTGTCTCCAGTGTCTATTTTCAAGTGTATTTTTCACTTTTCTTGAAGCAGCATGTCTGTT
 GCAAACTTCTAGAAATAATGAGAATATTTATATATTAGATCAAGCCATAACTTGATGAT
 ATAGTCATTTCTTCTTATATTTTTTACTTACATTTTTTACATTTTAATGATTACTTTTCATT
 TTTGAAAAACATGTCATGCTGAGATGTATTTTTCTTATTCTGTAATTAGTTATGAAACA

FIGURE 14.3

22/64

GTTTTTTCTTAAATGCTGAGTATATCAAGTCTTGGCTAAGAATAAGTAATAAATATTTGC
 CACATGAAAAGACTACACATATAGCCAGGTGCAGTGGCTTGCACCTGTTTTCCAGCTACC
 CAGGAGGCTGAGGCAGGAGGATTGCTTGAGCCCAGGGTTTCCAGGCTGCAGTGAACCTATG
 ATTGTACCACTCTACTCCAGAATGGGTGACAGAGCCAGGCCCCATCTCTCAAAACAGAAA
 AGAAAAGATTACATAGACTACATATACACCCCCATCCAAAACATACACACACATCTACTTA
 ACCTAAAATGGTAAGAAGATAACTTCTTATTTTCTAATATATGACACAGAAAAGTTTTTTT
 TAAAGTAGTTTTTAAATTTTTTAAATTTTTTCTAGGTATTTCTCAAGCCATGTTCCCATGTGG
TATCTTGTCAACAAGTTGAGGTGGAACCCCTCTCAGCAGATGATTGGGAGATACTGGTAA
 AGAAAACCAAATAAGAAGTATCTCATTTAAGGTTAAATTACTTCACAATATCAATGTCTT
 TAGCTTTCTCTAAGCTTTATTTATATATTTCTGAGTTGGTTTTTGAATTATAAGAATGAATTG
 GGGCCAGGCACAGTAGCTCATGCCATATAGTCCCAGCACTTTGGGAGGCCAAGGCAGGTGG
 ATTGCTTGAGTCCAGGAGTTCAAGACCAGGCTGGGCAACATGGTGAAACCCCGTATCTAC
 TAAAAATACAAAAATTAGCCAGGCATGGTAGTGATGCCATTAGTCCCAGTCACCTGGGGA
 GGCTGAGGCAGGAGAATCGCTTGAGCCCGTAAAGTCAAGGCTGCAGTGAGTCAGGATCTT
 GCCATTGTACTCCAGTCTGGAAAAACAGAGTGAGACCTTGCTCTCAAATAAAAAAAGAATGA
 ATTGATAGAGATCTAATGTACAACCTGACAACTATAGGTAATAAAATTTGATTGGGGATT
 CATGTTAAATGAGTAGATTTTTAACTACTCTTACCACAAAAACACAAAAGTGGGTAACTGT
 GAGATGATGTATATGTTAATTTACTTCACTATAGTAACCATTATCTATCTATATGTAGC
 TCATAACACCATGTGCTGTATATTAATATGCACATTAAATTTTGTTTTTTAAAAAAGA
 ATTTGAGATTTTTTTTAACTAGATATGGAGTGGACAAAATGTAAAGTGAATTGATCTTTTC
 GTCTGTTGGTTCTAGGAGCTGCATGCTGTTTCCCTTGAACAACATCTTCTAGATCAAATTT
CGAATAGTTTTTCCAAAAGCCATTTTTCCCTGTTTGGGTTGATCAACAAAACGTACATATTT
ATCCAAAATTGGTAGGTGCTATTGTAATATTTGCTGTCTATTTCTACACTATAGCATTGAG
TCCAAAGTAGAAATGAATGTGCACTAATGAGCTTTATTTTCTACACAGTTGCACTAATAC
CAGCTGCCTCTTATGGAAGGCTGGAACCTGACACCAAACCTCTTATTCAGCCAAAGACAC
GCCGAGCCAAAGAGAATACATTTTCAAAGCTGATGCTGAATATAAAAAACTTCATAGTT
ATGGAAGAGACCAGAAAGGAATGATGAAAGAACTTCAAACCAAGCAACTCAGTCAAATA
CTGTGGGAATCACTGAATCTAATGAAAACGAGTCAGAGATTCCAGTTGACTCATCATCAG
TAGCAAGTTTATGGACTATGATAGGAAGCATTTTTTCCCTTCAATCTGAGAAGAAACAAG
AGACATCTTGGGGTTTAACTGAAATCAATGCATTCAAAAATATGCAGTCAAAGGTTGTTT
CTCTAGACAATATTTTTCAGAGTATGCAAACTCTCAACCTCCTAGTATATATAACCGCTCAG
CAACCTCTGTTTTTTCATAAACTGTGCCATTTCATGTATTTCCATGGGACCAGGAATATT
TTGATGTAGAGCCCAGCTTTTACTGTGACATATGGAAAAGCTAGTTAAGCTACTTTTCTCCAA
AGCAACAGCAAAGTAAAAACAAAACAAAATGTGTTATCACCTGAAAAAGAGAAGCAGATGT
CAGAGCCACTAGATCAAAAAAAAATTAGGTTCAGATCATAATGAAGAAGATGAGAAGGCCCT
GTGTGCTACAAGTAGTCTGGAATGGACTTGAAGAATTGAACAATGCCATCAAATATACCA
AAAATGTAGAAGTTCTCCATCTTGGGAAAGTCTGGGTTAGTATAAAATTTTATAACTTGGG
 AGAAATTTTATGTGGCTTAAACATCCCCAAATTAATGAATTAGAATAGTATTTTCATATATA
 AATTGAAAATCAATTAATAAGAAACACAGTGCCTAAAGGCACCTTGGGGGACACATTTACG
 CTTTGCAGTAAAGTCCTTGTGTTGGATAAAGATTGTATGTTTTTCTGGCCAAGTAAGCTTGA
 ATAGGTACAAGCTTAGATAGGTTTCAGGCCAGAGAGGTCAAATTTACTTGCTGAGATTGC
 ATAGCTAGTGTTACAACCTAGGATTCAAACCCAGGCAGATTGACTTGGGGGTTTCATCAGGA
 TGGAGTGCCCTACAAAGCCTCCCATCTTTAATGCTTGACAGATTTGTTCCCCAGTTACCGA
 AAGCAACTTGTTAATATTAGGGAAAAGGGCCAGTGTAGGGAGAGATCCATGGCATGAGGT
 AACCTTCCTGCTGCATGTGGTGGCACCTGGATTGGAATGCATCCAGGAGCTGCTTACCCT
 GCCGGTGTCTGCTCTTTAATTTGTGTATAACGGAGAGGAAGTAGACAGGGCAACTAGTGC
 TCCAGCCCCCTCATCTGGCCACAAATATTAATGCTACCTTTATATGACATAAGTCACTAG
 TCCATTTATTGGAACCTAAATTTGAACCACTGTAAAGTAAGACTTCATAGTGATAAAGAG
 AGGAACCTGTTAGGAAAAGAGAATAAAAATAGAAAAGAGAAGGTTGTCTCCTTTTGTAGATTT
 TTTTTTTTTCTCCAACAGTTTTTACCTGTGACCTTTATACAAATAACTGACAAAGCATTA
 TCTCTTTGGCCTACATCATTTTCTTTTCTATTTTTTTTTTCCACAAGATGGAGTTTCACT
 CTTCTTGCCCAAGCTGGAGTGCAGTGGCATGATCTGGCTCACTGCAACCTCCGCCCTCCCA

FIGURE 14.4

23/64

CGTTCAAGTGGTTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGCATGCACCAC
 CACGCCTGGCTAATTTTTTTGTATTTTTTAGTAGAACTGGGTTTCACCATGTTAGCCAGCC
 TGGTCTGGAACCTCTGACCTCAGGTGATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGAT
 TACAGGCATGAGCCACTGCTCCTGGCCGGCCTACATCATTTTTCTAAAGCTCCAGACCATT
 CTTTTCTTTTTCTTTTTCTTTTTCTTTTTCTTTTTCTTTTTCTTTTTCTTTTTCTC
 TTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTTTTTCTTTTTCTTTTTTTGAG
 TTAGAAGCTTGCTTTGTTGCCCAGGCTGGAGTGCAGTGGCACCACCTCCACTCACTACAA
 CCTCCACCTCCCAGGTTCAAATGATTCTCCTGCCTCAGCCTTCAGAGTAGCTGGGACTAC
 AAGTGTGCGCCACCCTCCTGGCTAATTTTTGTATTTTTTAGTAGGGACGAGGTTTCACCA
 TGTGGGCCAGGCTAGTCTTGAACCTCCTGGTCTCAAGTGATCCGCCTGCCTCAGTCTCCCA
 AGGTGCTGGGATTACAGGCGTGAGCCACTGTGCCTGGCCTCAGATCATTATTTTTCTGTTA
 GCTTTAAACTGTCCGTTTCAGGAGATCCCCTGCATCCTCAAATTCAAAATATCTAACACT
 GAGCTTATGATTTAGCTGGTTCTGTCTATTAGATGGGAATATCCTTTTATTTCTTTGAAAT
 TATATGGTGAGAACAGGGAGAAGTGCTGATGGTAAAGTCTGTGATTAAGATAGCAATAA
 GGACTCCGCCCTTCCCCTCCACTGAAGGTTGAAGAGCCATGGACAATGAGAAGTCACAG
 TAGGTGAAATCAGGTACTAAAATGGACTTGGCTTGAGAGATCAAAATTGATCACTTGGTG
 ATACAATAACAAATTCATGTAACTTGAACCTTTATTACCCTGTGAAGCATGGTGATTA
 AAAAAAACAACAACAACAGGAACTTGATTGTTAAATTCTCTTTAAGTCAGAATATG
 TACCTTAGAGTTTTTATTATGCTTTTGTCTTACCATTAAATATGTCTGCACCTGCTCTTTA
 GAAGTTAATAGAGAGTAAAGTCGTCTTTATGTCTTTTCAGTGCTTACTTATATTTGGGAAG
 TTGAGAAAAATTTTTAACATCATTATTGATATATATATATATATATATATATATATATAT
 ATATATATATATATATATATATATAGATAATTTTTTTTTTTTTCTTGAGACGGAGTCTCACT
 CTGTGCGCCAGGCCGGAGTGTGGTGGCGATCTCCACTCAATGCAAGCTCTGCCTCCAGG
 TTCAAGCGATTCTCTTGCTCAGCCTCCCGAGTAGCTAGGATACAGGCTCCCACCACCAC
 GCCTGGCTAATTTTTGTAGTTTTAGTAGAGACGAGGTTTCACCATATTGGCCACGCTGGT
 CTCAAACCTCTGACCTTGTGATCCGCCCACCTCGGCCTCCCAAAGTGCTGGGATTACAGG
 CGTGAGCCACTGCGCCCGGCTGAGGTAAAATTTAAAGTGTAACAATTCAGTCATTTTTAGT
 ATATTTATACTAGTTGTACAGCCATCACCACAATCTAAGTTTAGAACATTTTCATTAGGG
 GGTGGGAGAAATTTTACTCTGCTTTTTAGATTAAGTTTCTGTCTGGATCTAATCATTAA
 TCAGACAATCAGGCAGATTGTCTGTGATTAGTTTTGGCCATTCCAGCTTCTTCATTGGTT
 GTTAACTTTCAAAATAAAGGCTGCTCAAAGATTAGAAATAACATTTAATTTGAATGTAA
 ATGTGCCATAGTTTTAAAGATGGGTTTGGTGAATACAGTCAAATACATACATTTAAAGCT
 CTAATTTCTGAAGATTATGTAAAGAAAAGGAAAGAAATGTAGGGAGAGGATTGAAATGTTT
 ATGGTATAACAATATCTGAACATCCATCTGGTCACACCGTTGGTATTTGAATGTTTTGTC
 CTCCTCAAATTCATATGTGCAAAATCCCACTCCCAAGGTGATCGTATTAGGAGGTGTGGT
 CTTTGGGAAGTGATTAGGTTCATGAAGGTGAAGCCTTCATGAATGGGATTCGTGCTCTTAT
 AAAAGAGAACTGTGAGAAATAAGTTTCTGTCTGTTGTTAGCCACCCAGTTTAGGATATTT
 TGATATAGCAGCCTGCATGGACTGAGACAACATAGAGTTATTATGATAGCTTCTGTTATT
 TCACCTAAATTCATAGAAGCTAATATATCAATATTTATGCTATGAAATATTTCTTAACCA
 AGCTTTGAATATATTTATATTTTTGTTTATTTTTAAATTTTCAGATTCCAGATGACCTGAG
GAAGAGACTAAATATAGAAATGCATGCCGTAGTCAGGATAACTCCAGTGGAAGTTACCCC
TAAAATTTCAAGATCTCTAAAGTTACAACCTAGAGAGAATTTAGTGAGTTCAAATATATA
 TGTTACATCAAATTTCTTTTACACGTTTTGTAAGATTTCTAGTTGCTTTAGCTAAGTAAT
 AAGAATGTTGTATTCTTTTTGATACAAATCTTTTTTTATTGTGTAACTATATATAAC
 ATAAATATGCCATGTTTCGCCATTTTTAAGTGATAATTCAAAGGCATTAATTACATTCA
 TAATATTGTACAACCATCACCCTATCTATATCCAGAACTTTCCATCACCCCAAAGAGA
 AACTTGGTACCCATTAACAATAATCCCCGTCCACTCCTTTCCCCAGTCCCTGGTAATC
 TCTAATGTATATTGTGTCTCTATGAATTTACTTATTCTAGATATTTCTATATATAAGTAGA
 AGTATGCATTTGTCTTATGTATCTGACTTATTTCAATTAACATAATGTTTTCAAGGCTCA
 TCTGTGTTGTATGTATCAGAATGTTATTCCTTTTCATGGCTGAATACTATTCCATTGACT
 GCATATACCACATTTGTTTATCCATTCATCTGTTGATGGACACTTGGGTTGTTTCCACAT

FIGURE 14.5

24/64

TTTTGGCTGCTGTGAATAATGCTACAGTGAACATTGGTGTACAAGTATCTGTTTGAGTTC
 CTCTTTTTCAGCTCCTTTGGGATATACCTAGGAATTATGTTTAACTTTTTGAGAAGCTGAG
 AAATCTTTAATAAATGATAACACAAATACTTATATTTGCCAATGCAAATATGAATATTTT
 TGGCTTTTAAGAGATTGATCATTTTGCCACGTGGTTGTAATTAATAAAAAAATTGTCCCATG
 TTGTTTCAGTATTAATATTGTAGCCTAAAAGAGTGCTAGACTGTTTTACTTTTTACTCAG
 TTAATTCCTTTGGATACTGGTAGAGTCAGGAAATGAGATATTGAACTTAAAGATCTTTGCA
 GGTGGGGTCCAGTGGCTCACACCTGTAATCCTAGCACTTTGGGAAGCTGAGGTGGGAGGA
 TTGCTTGAGGCCAAGAGTTTGAGAATAGCCTGGGCAACATAGCAAGACCCCATCTCTACA
 AAAAAATTAAAAAATAAAGCCAGGCGTGGTAGCTCACGCCTGTTATCCCAACACTT
 CGGGAGGCTGAGATGGGTGGATCACTTGAGGTGAGGAGTTGGAGACCAGCCTGGCCAACA
 TGGTGAAACCCCATCTCTACTAAAAATACCAAAATTATCGGGGCGTGGTGCTAATCCTGT
 AATCTCAGCTACTCAGGAGGCTGAGGCAGGAGAACCCTTGAAGTGAAGGAGGTGGAAGTT
 GCAGTGAGCCTAGATCTCACCCTGCACTCCAGCCTGGGTAACAGAGCGAGACTCTATTT
 CAAAAAAGTAAAAATAAAAAATTAGACACATGTGGTGGCACATGCCTGTAGTCTTAGCTA
 CTCAGGAGGCTGACTGAAGTGGGAGGATCTCTTGAGCCCAGGAGTTCCCACTGCAGTGA
 GCTATGATTGTGCCACTGCACTCCAGCCTAGGCAATATCTCAAAAAAATTTTTTAAAT
 AGATTATTAGGCCAGACGTGGTGGCTCATGCCAGTAATCCAGCACTTTGGAAGGCCAAG
 GCAGGCGGATCACCTGAGGCCAGGAGTTTGAGACCAGCCTGGCCAACATGGTGAAACCC
 ATGTCTACCAAAAAATACAAAAATTAGCTGCAATGTCTATAATCCAGCTACTTGGGAGCC
 TGAGGCAAGCGAATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGTGGAGACTGCGCC
 ACTGCACTCCAGCCTGGGCGATACAGCGAGATTCTGTCTCAAAGAAAAAGGAATTTGTTT
 TCCTGTCTTTATCGTAGAGGGAGGAAAGGAGAATGGGGTTGGAATGGTTATTGAGTGAG
 CCACATTATGGTAGATGTATCACTGGGCATAGAGAAAAGGAGCATTTAAAACTTTTCCGC
 CTAACAGATGTTTCTTCAGGCTACACTGCACTCATTGTGCTAACTGTAATGTCAAATCCC
 AGACCTGTGCCTATAGAACATGAACATCCTTCATTGGATTGTGTTGGTCAGGCTTACACT
 TTATTAGGAAGATCAGATGTTAAAATAAGGGTGTTAAAGTTAAGTTCAGATATGAGGATA
 ATTCATTACTATTCTTTTTCTGGCAGCCTAAAGACATAAGTGAAGAAGACATAAAAACT
GTATTTTATTTCATGGCTACAGCAGTCTACTACCACCATGCTTCCTTTGGTAATATCAGAG
GAAGAATTTATTAAAGCTGGAACTAAAGATGGTGAGTACATTTGTTATTTTGAATTTTTT
 TTCTATTTAAATAGTTGTACATTTTAAATTGTTCTTGCAACCTGTCATACCTGTGAACAG
 TATGTGAATAGTGAAATATAATTATGATAATTAAACAGTAGTTTTTATGTATTGAAAAAT
 ATCTTTGGCCGGGTGCAGTGGCTCATGCCTGTAATCCAGCACTTTGGGAGGCCGAGGCA
 GGCGGATCACTTGAGGCCAGGAGTTCGAGAGCAGCCTGCCAACATGGCGCAACCCATCT
 ATACAAAAAATACAAAAATTAGCCTGACATAGTGGTGTATGCCTGTAGTCCCAGCTACT
 TGGGAGGCTGAGGCAGAAGGATCACTTGAGCCCAGGAGGTCTGTGTTCTGCCACTGCAC
 TCCAGCCTGGGCAGCAGAGTGAGACCCTGTTGGGGGGAAAAAAGTCTTTAACTT
 AAATAAATTTGACATTTAAATCTTAAATTATTTTCTCTCTGTTTCAGTACTAACTCTGC
 ATTTATTACTTTCTTTTAAATAGGACTGAAGGAATTTTCTCTGAGTATAGTTTCATTCTTG
GGAAAAAGAAAAAGATAAAAAATTTTTTCTGTTGAGTCCCAATTTGCTGCAGAAGACTAC
AATACAAGTAATAGCATGTTATTGAATATTTAATAAAATACTATTTGTTACATATGATTG
 ATAATAAAGTATGAAGTTCCTTGTAACACCTTGCAATTGTGAAGTGTATTAAAAACCTGCT
 AAGAGTAAGGAATAACTTGATTTAAATATTTTATTCTGTAATCTCTTTAAATTATCTGT
 ACAAATTATTGACTTAACCTAAATTTAAAAATGAATGCCTTAGCACAAATTAAGTTCCAAG
 AATAGAGTTGATCATGTAACTGGTAAATGGATCATGATTTAAATCTTCTAGGATTGA
 AACAAATGAAAACGTAGTTTTTAAGGGTTTGATTTTTTAAATTCCTATTTTTTACATGCAAT
 TTTACTGCACAACCCATCTTATTTTGACAGTTCCTTAAATTCGCAACTCTTCAGAAATATT
 ATCAGATCACTTTTCTTTGCTTCCATAAGTTTTTTTATTATTATATTATTTTTTTTTT
 TTTAAAAGACGGTGTCTCACTTTGTGCGCCAGGCTGGAGTGCAGTGGCATGATCATGGCT
 CACTGCAGCCTCGACCTCCCAGGCTCAGGTGATTCTCCACCTCAGCCTCCCAAGTAGCT
 GGGACCACAGGCGAATGCCATGATGCCTGGCTAATTTTTGTATGTTTTGTAGAGATAGGG
 TTTTACCATGTTGCCCAGAATTGTCTTGAACCTCCTGGGTTCAAGCAGTTGTTCTGCCTTG
 CCCACCCAAAGTTGTGGGATTACAAGTGTGAGCCACTGCGCCAGCTATTCTAGAAGTAT

FIGURE 14.6

25/64

TTTAAGAGTCATCTTTTTTTTTTTTTTTGAGATGGAGTCTCAGTCTGTACCCAGGCTGGA
 GTGCAGTGGCACACTCTCGGCTCACTGCAACCTCCACCTCCTGGGTTCAGTGATTCTCC
 TGCCCTCAGCTTCCCTAGTAGCTAGGATTACAGGCGCATGCCACCATGCCCTGCTATTTTT
 TGTAGTTTTAGTAGAGACGAGATTTACCATTGTTGGCCAGGCTGCTCTTGAACCTCCTGAC
 CTCAAGTGATCTGCCCTCCTCAGCCTCCCAAAGTGCTGGGATTCTAAGTGTAACCACCA
 CACCAGCCAAGAGTGGTCTTTTTTACAATATTATTTTTTTGATTAGGACATTCAATCTTGT
 CATAAAATTGAAGATACTCTAGTCATTTAGAATTTCAATTGTTTTGGAACCTAGACATTGTT
 TCTTTATTTTTGAAATGTTATTGAAGGAATACCATTGGAGAAGATACAAATGTAAGAAT
 TGTGAAAAGGATAAATTGTGACACAAATCAAATTTATAGATAAAATATACCTGTAAATG
 TATTAAGGCAATAACATTCTTTCTGCTTGTGACCATAAATATTTATATTCCCTGGATGG
 GTACATTGTTATTTGTCAAGGGTGTAAATAATGATCTTGCATGCATAATTTATTCCTC
 TGGTATAACAGAATCAGCAATTTAGTTTTCTGGGACCCGAGAAAAACATGCAAAAGACAT
 ACTTTGAAATGTAAACTGATTTTTCTTGCACCTGTAGGTCCTTCTAGATCCTATGGTA
AAAGAAGAAAAACAGTGAGGAAATTGACTTTATTCTTCTTTTTTAAAGCTGAGCTCTTTG
GGGTAAGAAGTTATGGCCAACTAGCATGTTAGACATGTTTTTAACACTATATCTGGCAG
 AGTTTTCAATGTAAATATTAAAGTAGATGTTAATGTCAATAAGTGATCTTAATAATGCAT
 CAGTAGATATTTTTCAAGGATTGTCTCTATCTTCACGCCCTAGCTTATAATTTGCCCTTGT
 CGTCTTTTTTTTTTCTCTTTATTTTTATGTTTTATCCATCCCTGGTGGTAGGGGATAA
 CCTTGCTCTCTTCGATAACAAGAAGTCTGAAGCTTATTAGAAATTTTACTTTGAGAATTG
 ATCGATGAGAAGAAAGCAACTAGATATCACGTGGATCATATATGCTTGAATAAAACAATA
 ATTCCTTAGAACAAATAAAATACATTTTAAAGTTAAAGCCAAAAACATTAGTTGAATGTTT
 AAAAAATATTTCAAATTAAGTTATTCCTTCACTGTCTTGTATTACTGTAATAATTTGGATT
 ATTTGTGTTTTTCTCAACTTTTTAAAACAAATATTTAAAAAATTCCTCTTTGATTAAAGTA
 GGGCTAGATAAAATATAAAAAATATTTTTTAAACTCCTCTTAATTTCCATATTTCTTATA
 TAATATGAGAATCTCTTATAAACTACCTCTTAGAAGTCTCCACAGAAGCTTTGGTAGA
 TGTAGTAGTAGGGATTGATTTCTTAGAATGGTATAATCTGTAAATGTTTTAGTAAAAGG
 ATTAACGATAAAGTCAAAATGTTTATAGCACAGTGTTTATTAATATAAAATAAAATCTC
 TTTTTTTTTTTTTGAGATGGACTCTCACTTTGTCACTCAGGCTGGAGTGCAAGTGTGCAA
 TCTCAGCTCATTGCAACCTCCGCCTCCTGGGTTCAGCAATCCTTCCGCATCAGCCTCCT
 AAGTAGCTGGGATTACAAGCATGCACCACACACCTGCCTAATTTTTTGTATTTTTTAGTA
 GAGATGGGGTTTTACCATTGTTGGCCAGGCTGGTCTCAAGTGATCCGCCTGCCTCAGCCTC
 CCAAAGTGCTGGGATTACAGGCGTGAACCACTGTGCCAGCATAAAGTAAAAATCTCTCA
 GACTCTCATGTGATCATGTAAAGTGGCAGGCAGTCAAGTCAAGAGTAGTTTAAAGTTC
 ATGTTTGTAAATATAATCTACAGATTGATACTGGATTTCATAGGTAATGTTTAAAGAGAA
 AATAAGTTTTTAGTTATCCTCAGTACTTCAAAAGCACCCATTTATGATTATGTTGATTAC
 TAACTAAATCATTTGGGGGCTAGAGGTGTTTTTTTTATGTGTTAAGATTCCCTAAGGAGT
 TCTATTAGGGCAAACTTTTAGTAAGTGCATATTTTAAAGTAATAAACTAATTTTTAAA
 AGCTTGGAGGCTGGGCGCGGTGGCTCACACCTGTAATTCAGCACTTTGGGAGGCCAAGG
 CGGGTGGATCACTTGAGGTGAGGAGTTTGAGACGAGCCTGAGCAACATGGTGAAACCTTG
 TCTCTACTAAAAATACAGAAATTAGCCAGGTGTTGGTGGTGGGCACCTGTAATCCCAGCTA
 CTCGGGAGGCTAAGGCAGGAGAATTGCTCGAACTTGGGAGGCAGAGGTTGCAGTGAGCCG
 AGATCATGCCACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCGTCTCAAAAAAAAAA
 AAAAAAAAAAGCTTGAAAGTCAGATTTCGACATTAATCAGTATACTTTCTCTCAAGTAGGGG
 ACAATTTCTAAGATTTTAGTCTTTTTAAAAATTTATTAAGTCTGAGCATGGTGGCTTGT
 GTCTATAATCCCAGCACTTTGTGGGGCCGAGGCAGATGGATCACTTGAGCCCAGGAGTTG
 GAGACTAGCCTGGGCAACATGGCAAAACCCCGTCTCTACAACAAATGCACACACAAAAA
 CCAATCAGCTGGGTGTGGTGTACACTCCTGAAGTCCCAGCTACTCGGGAGGCTGAGGC
 AGGAGGATCACCTTTGCCAGGGCGTTTGAGGCTGCAGGGAGCTGGGTTACACCACTGCG
 CTCCAGCCTGGATGACACAGCAAGCCCCCTTTCTCAAAAAAAAAAAGATAAAAAATTAAT
 TAAATTAATTAAGTACACTGGGAAGGCAAAATTCAGCATTTTTTTTATAGCTAAATTTTAT
 CCTGCTTCAGTCTTTTATCATGTAACATATGTATATTTTTTACAGAGGAGTGAATTCCTTA
GGCGTATCCTCCTTGGAGCACATCACTCACAGCCTCCTGGGACGCCCTTTGTCTCGGCAG

FIGURE 14.7

26/64

CTGATGTCTCTTGTTCAGGACTTAGGAATGGAGCTCTTTTACTCACAGGAGGAAAGGTA
AGTGGTTAAGGTGTGTTTCATTTTTCTGTAACTTTAATAACTTTTCATTTATCTTTCTTT
GGGTTTTGACCATCTATTATATAGGGTGGGTTTTTGACCATCTATTATATAGGGTTTATAC
GACATATGGAAAGCATTTCATTTATTCACATAATTTCTGTGTGTCTGCTTTTAGGTGTTG
GGGGAGTGATGACGAATAAGACTGATGTTCTCCATGCCCTTTTCTGTGTCAGTTGATAC
AATTATATGGTTTTCTTTTTTAGGCTATTAGGTGTTGATAGGGTTGAGTAACCTTACAAA
TGTTGAACCAGCCTTGCATACCTGTGATAAATACCACGTAGTTGTGGTGTATCATTCTTT
CTACATTGCTGAGTTTTATCTGCTAATGTTCTGTTGAGCTTTTGTCCATTTAAGTTTGAA
AGTGATTAGTTTGCAGTTTTCTGTTTTTGTGTTGCTTTTGTCTGGTTTTGCTATCCGTGT
AAATCTGGCCTCATAAAATGAGATGGGAAGTATTCTCTCCTCTTCTTTGTTTTTTTGA
AGAGTTGTATAAAATGAGGCTGAATCTTGGTGGTTGCCACAATGACAGGAACTATTTTC
TGTGACTGAATATATTGGGAATTCCTATAAAGCAATTATTTCTAGGGAAGTGGAAAAATC
AACCTTAGCCAAAGCAATCTGTAAAGAACATTTGACAAACTGGATGCCCATGTGGAGAG
AGTTGACTGTAAAGCTTTACGAGGTATGAGTATGGTAACACTCTATATAAATCCCTTTTT
CATTAGAAAGACAGGAATGTTATACATAATGCTGTCAATCTAATAAATACACATATCATC
TAGTCTTTAACTTTTCTGTTTATCATTTAGTCATTAAAAATTTCTTTGGCTTTCTAATGTT
TTTGATAAAATTTCTAAAACTCTCCATATTTAATGGAGGCCTATTTTTTTTTCTAGCCAG
AACTTTTTGTAGACTACATTTCTGGAAGTGCTCACTGACACCACTCTGAAAAATTAGTAC
TTAGAATATACTCTAATTGGTATAAATGATCTCTGAATTGCTATGGAAAACTGGGAGAAT
GGTTGCTTCAGGGGAGAGAAAGTAGGAGGCTGTGGACAGCAATGAGGAGAATTACAGTTC
ACCATATAACACTTTTGTACTTTTAAAGTCCTTAACATTTACATTATTATCTATTCAATT
AAAAAATATTGGGAAGATTTTACTTTGAACAGTTAATTTTTTCCCCCATGGGTACCGCTGT
CATATAGTTCCAACTAATCATGAACCTTGTGTATTTTCTGTTCTTTGTAAATTTAAACTTT
GTAACCTCACCAGGAAGTTTGAAGCCAAATTTGTGTTTCAAATATAGCAACTCCAGGATCT
CTAGGCAGATGCATTTGCATTTGATTTTAAATGAATCTTGATCCCTTACTCTCACTTATG
TTTTCCACATCCTACTTTTTTTTATTTTGTGTTAAGCCATCTAAAAATCTCAATGGGATG
AAACTGGGTATAAATGAATACATGCATACAGGAATTATAGTAGCATATTCCTTTCTTTT
TTCTTTTTTTTTTTTTTTTGAGACAGAGTCTTGCTCTGTAGCCCAGGCTGGAGTGCAGTGG
TGCGATCTCGGCTCACTATAGCCTCCACCTCCAGGTTCAAGCAATCTCGTGCCTCAAC
CTCCCGAGTAATTGGGACTACAGGTGCATGCCACCACACCTGGCTAATTTTTGTATTTTT
TAGTAGAGATGGGGTTTACCATGTTGGCCAGGCTGATCTCAAACTCCTGACCTCAAAGT
GATCTGCCTGCCTTGGTTTCCCAAAGTGCTGGGATTACTAGCATAAGCCACTGCACCTGG
CCTCCTTTTCTGAGTTTTTATAAAATTTGATACTTTTACTGCACGCTTTGAGACTGTATTAA
TTGAACCATGTTGATGAACAAGTTTTTGTGATGGGTATATTAATAAAATATAGATCAAAT
TTTTATAGTTAAATCAATATCGAGCTTTTCTAGTGCTTTCAAAGGACAACCTGAATTTT
CCCAGCACTGAAATGATACTGAAACCATTTTCATATCTTCTGTATTAAGGAAAAAGGCTTG
AAAACATACAAAAAACCTAGAGGTGGCTTTCTCAGAGGCAGTGTGGATGCAGCCATCTG
TTGTCCTGCTGGATGACCTTGACCTCATTGCTGGACTGCCTGCTGTCCCGAACATGAGC
ACAGTCCTGATGCGGTGCAGAGCCAGCGCTTGCTCATGTTAAATGCATCCACCCTGGC
TTAAGGTCTTGTCTTTTGTTCAGTCAGCATTTTGTAGTCTTAACAATAAATCTACTCTTT
CAGAGAATAATATATAGTGTATGTTAAGTGTGTGTTTGGAGCCCCCTGATGGCATTCTAC
AGTTGTCTATAGACTGTAATAGCAAAATTTGGTAGAGTAAAAACAGTGTGAAAAATCTGC
AACTTCATGGTTAGTCCCTTTAGGGTTTTTTCATTCTCCCTTACTTATTGTTTAAATTTACAG
ATTTACTCTTTTGTTCATTTGACAAATATTTGTCAAATGCTTGTGCACAGTCTGTATTCT
CAAATCTAGGAGAAAAAGAGGGTGAACAGTATTAGCGCAGAACGATACTAATAATGAT
GGCTACTGTGTATGAGTAGCCAGCCCTTTCTTGGCTTTCTTGGATTGCTTTGTATTCTAC
ATGAAGATATTCCCTGGGCTTTACAGGTCAATAAATGGAAATTCAGAGAGATTAAATTTGA
CCAGGGTGACCAACAAGGAGATGACAGCATACTATGCGAGAAGTATACACAGAGTAGT
GTAGGAGCATATAACCTAACTGGGGGTGAGGTGGGATAAGGAGTTATCAGGGAAGGCTT
TTTGAGGAGATTGACAACTGAGCCGAGTTTTGTGGAAGAGTAGAAATTAGCATGAACCA
ATTTCATGCTAATAAAGAAGCAAAGGAAGCGTGGTCTACAGGCAAAAGCACAGAGGTACA
GGAAGTAATGATATGTTGGGAATACCTGTTGACTGGAGCTTAGAGTGCAAGGAGAGGA

FIGURE 14.8

27/64

GTGCTAGGGAGGTGAGGTTGGAGGGTTTGGCAGCATTGACTTGCTTCAAGGTTCTTAAGA
GCTGAAATAGATATAAAATGCAACTAAGAGTGGCTTGGATTATTATTACCTAGTGTGTTA
ATCTCAAATTTTGAAATCTATAGCATCTATAGGACTGGTGTTACTAATCTTACACTCGAT
CTGTTACTGTTCTTATACTAGATCTATTAGTCCAGTGTTTAAGGGAGTGGTGCAGATTTC
TAGGTCAGGACAGGACTCAGATGTACATTATTAATGCCTATTTTCAGTTCTGACCTTCTCA
TATGAAACCTTATAAGACCTGGGGTAGGAAGAGATTGTTCTGGAAGTCATAGGAATATGA
ACTGTATTTTGTTTAAACAAACAATACAGTATGGAAATTTATCACCTTCCAGAATATTTA
TTTCAGAGACAAATTTTATCATTTCGTTTCATTATTTTCATAAGATCCACGAGTAGGGAAC
CTCACTAGACATTGCTCTGAGTATATGGTCTGAGTTTGCAGTACCTCTTGTGTCTCCATT
AGATTTATTAGGTCCTCAATAGATAAATCAGGGAATAACTAGATGGATTCAATTTTAAA
GACATGAAAGAGCGATACCATACATACTGCACCTTAAAGGTCAACCTTAGAGTATCATTA
TTTTTAATGAATGTATAATTTTTTAAATTTTCATGTTTACTTTTCTTAAGCTTTTGCACAT
ATTGCTTAATTCCAGCTTTGAATGATATGATAAAAAGAGTTTATCTCCATGGGAAGTTTGG
TTGCACTGATTGCCACAAGTCAGTCTCAGCAATCTCTACATCCTTTACTTGTTTTCTGCTC
AAGGAGTTACATATTTTCAGTGGCTCCAACACATTTCAGCCTCCTAATCAGGTAATACACT
ACTTGTAAGGATTATTGAATTATGTCCCTTTTATAGAAATTATTTTTCAATTTTATTAGT
AATTCGTGGCTTTAAATTTATGCTTCTCTTAATGATTTTAAGGATATGTAAGTCAACATT
TGGTGCATATTGTGCTAGAGGCATAAATTATAATTTATAGCCACCTGAAATGTTAGTATG
CGCTTTCCAAGAAAATGACTTTTTTGAAGTGGTATTCTTTGAATGAGAAAGAACAGAG
AGAAATAGATAGATGGCTTTTAAACACTTCATTAATTAACCTTTTTTTTTTCCACCATCAC
ATAATGGCACTTAGTCCCCCTTTGGGAACCTCATGAGGGTTTATGTTGGTAGTGAGCTGAAAG
AAATATGTTCCAGGACTGGCAAACATATTCTAAATTCCTTTAAATTTTTCACCTAGCATCT
ACCCTAAATATTTCAGACCCTGTGCTAGTTAACTGCTATTGAAGAAACAAAGGTATTATATC
TATTATTAAGGATAATAGAATGGTATTTGAGATATTGGTCATTGAATATGAATATGTTTT
GAGAAATAAGTTTTATAGGAACCAAAAAAATTCTTAAAGGAACCATATATTACTAAAA
ATGCTTCTTATTGGAGAAAGAAATGACAATCATTTATTAATGTGATTTTTTACAACTTT
ATTAAGATATAATTTAAGTACAACAACTCACATAAAGTGACAATTTGATCAGTTTTAA
CATATGTAGATGCCATGAAACCATCACCACAATTAAGGAAACAAACATTTTCATCACTCC
AGAAGTCTCCTAGCCCTTTTACTACCCATTCTCCCTGCTCCATCCCCAGACAACCTACC
AATTTGCTTTCTGTCACTATAGATTTGTCAACCTGATTTTCTCCAAATATACATTCAAAA
ATATACAGTTGAATACAATTGGAAATTCGAATTTTGTGTTTTTTCTTTAGGAACAAAGA
TGTGAAATCTGTGTAATGTAATAAAAAATAAATTGGACTGTGATATAAACAAGTTCACC
GATCTTGACCTGCAGCATGTAGCTAAAGAACTGGCGGGTTTGTGGCTAGAGATTTTACA
GTACTTGTGGATCGAGCCATACATTCTCGACTCTCTCGTCAGAGTATATCCACCAGAGAA
AGTATGTTTTACTATTAACCTGAACTTGAATCTTCTTTCTATTGTGGAGAAATGTAA
TTGTAGTAAGACAAGAATTAAATATATTCCATTGTAGTATTTGAATAAGCAGTTATTTGA
GTAGAAAATTAGTGTTCAGCTAAGATGATGGCATATTTTGAAGTTCATATAGTGAAT
ATACTAGTAAAAGAAGTTTGTATTATTTTAAACAGAAATTAGTTTTTAAACACATTGGAC
TTCCAAAAGGCTCTCCGCGGATTTCTTCTGCGTCTTTGCGAAGTGCAACCTGCATAAA
CCTAGAGACCTGGGTTGGGACAAGATTGGTGGGTTACATGAAGTTAGGCAGATACTCATG
GATACTATCCAGTTACCTGCCAAGGTATGTTTAAAAAAGAAAAAGTGAATACTTACTCC
CAGAAGAACCCTGTATTATTGGCTTTGGCTTTATGTGTCAGCTTGCCCAATCTCCGTGT
GAGTCAACAAGTGTTTACTGAGTTACCAATAAATGTCTTAACACTATTTTAGGTACTTT
AACAAATTTTAATTTTATTAATTAATTTTTTATTAGAATTGAGACCTCACTCTGTCTCT
AGGCTGGAGTACACTCACAGCTCACTGCAACCTCAAACCTCTGGGCTCAAGCAATCCTCC
TGCCTCAGCCTCCCCAGTAGCTAGAACTACAGGCATGAACCACCATGCCCGGCCAACTCT
TTAATTTTCTTAGAGACGGAGTCTTGCTATGTTGCCAGGCAGACAGATTTTAATGTGTA
TGATGCAGTCTTTGATGATAAGAACTTATAATGGAAAGCTGAGGTGATAGTTACAGTAA
ATACATTTTGATGTATAATTCTGTTTGCTTTAATCATTCAAATTGTAGTAAAGCAAGATG
AACTGTCTGCTGGGATTTGAGCAGAAATGGATAGGAATAAACTAGGAGGTAGAAGAGTTA
TCAAGGTTTCAGGACTGATGGGTGAAGCTAGATTTCCAGACCCGGGATGTCAAGTCTTG
AAAAGCAGACTTGGCAGGCATAGACGAGGCAGATAGCAGGATAAAGGAGACAAATGTAGA

FIGURE 14.9

28/64

TTGTTCTTCAGAAGATCAGATGGTAGAGTCTAGGAGGTAGTGTGTTTAAATCAGAGATCT
 GAGAGGCCAAAGATCATTGCGATGAGATCAGGGACCCATGCAAAGGAGTGAGAAAAAACT
 GGGTTAAGGAGCCTGCTGCGATGGCAACTCCTGGGAACAGTGGCCACTGGGGCCTGGGACA
 TGTTGATTGCAGCCAGGACTGTTAAAACAGTGTGAGAGAACATGGGTATGGAAGTACT
 AGCTAGCAGGATCATGACCCGATGCTGGGATGGGGCATCAAGCATTAGTACATGGAGAT
 TCAGTACATCCAGATGCAGTACATGGAGACTATATGCGTAACTGCTGACTTTGGGCTTCT
 TTCAGATTGGAGCAGAGGTAGAGGTGAGTGGGAATATTCTCAATAGAGGGAATAAATAG
 GCATACCTAATAAAGGAGACCAGGATATTGCAGACAGTAGCCTCATGTTTGGCTCACCTG
 TTCAAAAAGTTCTCTTGTCTTGTAGCAGTGGTGCCTTAAAAGGTAACCTTGAGAAGCAGTC
 GATTATTTGTTTCAGCCTGGAGACTCTTGGGATATTTTACTATCTTTGATTGAATAGATTT
 AAATGTACACAGCTCTCATAACTTGCCCCATGAAGCATATCCATGAAAGGCACTATACTT
 GTTAAAAGATTGGTTTGTACTTTTTTAAATGTAGTACTTTTAAATAAACAGGAAAAATAGA
 AGTTCTGATGCAGTTATATGCATTTTATATAGAATGTGTTCTTAATTGGAAAAAATTTGT
 CGTAGTTCCTTTGAGTTCATTTACAGTTTTTTAGTAGGAATTGTATTTTCTACTGTTGTAC
 TTGCTGTTACTAAAGAAAGATGGTCGTGATTACCATCTGAATTTTTTTTCTATACATTGA
 TCTTTAGCTGCTACTTAGTCATTTCTGTTTAGACTTGAGCTCTTTTTTCATATTTTTTTT
 TTTGTTTCTCAGTATCCAGAATTATTTGCAAACCTTGCCCATACGACAAAGAACAGGAATA
CTGTTGTATGGTCCGCCTGGAACAGGAAAAACCTTACTAGCTGGGGTAATTGCACGAGAG
AGTAGAATGAATTTTATAAGTGTCAAGGTATGTTGTCTACTTATCTTCTTTTTTATTTA
 GGTAAAATTAACATAAATGCAGTTAGCCATTTCAAAGTGTAATTCAGTGGCATTTAGTG
 CATTACAAATGCTATGCAACACCACCTCTCTCTAATTTCAAACCTTTTTCATTCCACTC
 CTCCTCTTGTCTATCCCCCTGGCAACCATTTCATCTGCTTTTTGTCTCTATGGATTGCTT
 TTCTGTATATTTTATATAAAACAAATCATGCAATATGTGACCTTTTTTGTCTGGCTTCTT
 TCACTTATGTAATGTTTTTCATGGTTCATCCAGGTAGTAGCATGTATCAGTACTTCATTCC
 TTTGCATGACTGAATAATGTTACCATACTTTGTTTATCCACTTATCAGTGGTGAACATTT
 GAATTGTTTCTACCTTTTGTACTATTATGAATAATGTTGCTGTAAATATTCATGCACAAAT
 TTCTCCACGGATATGTTTTTCACTTCTCTTGGGTATAAACTGAGGAGTAGAATTCTTGGGT
 CTTAGGGTAATTCTCTAACTTTTCAAAGAACCACCAAACCTGTCTTTCACACCAACTGCAC
 CATTCCCCTAGCAGTGTGGGGGGTCTGATTCTCCACATCTTTACCAACACCATTATG
 TTTCTCAATTGTGGGCTAGTCTCACATTTGGAAAGCTAGTGGGAGCAGCGATCCATCTAT
 TAAAAGTTGTATGAAATTGAGTAATGAGCCACCTCTCTCTTGTAGGGCTTATTATGTTCT
 TGCTTAAGGCAATCTTCATGCATTGTGAACAGAATTATACATAAATGCTCAGATAAAAGG
 GCAAACCATTCTTAAAGGGAGTAGACAACCTAGAGGCAGGAGACCATACTGAGGCAGGAAG
 CTGGGGTTTTTATGGTTCTGTACTTTTGTACTATATCTCACCATTGCTTTTGTCAAAGTG
 AGACTAGGTCTAAGTTTTTTTTCAGGTATAAGGTGAGTGTGGTAATTAAGGGGCATGCTAG
 CAGATCATTTTGGGTAATGCTTCACAGTCCACCCTGGTGTGTCTATTGTGGTTCGAGATC
 CAGTATCTTAGCTGTGTAATTTAGACATCAGCAATATTAGTTTAAACAAAGGGCAATTAG
 ATTCCAAGACAAAGGAATCGTGTATTATTCTAGCCTTATTCAAACCTGATTATATAAATCA
 GTTTAGTAATTTATTTATTTGTTTCTGTATTATTATTTTATTCTTTGAGATGGAGTCTCA
 CTCTATTGGCCAGGCTGGAGTGTAGTGATGCAATCTTGGCTTACTGCAACCTCTGCCTCC
 TGGGTTCAAGCTATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCTAATTTT
 TGTATTTTTTAGTAGAGATGGGGTTTACCATTGTTGGCCAGGCTGGTCTTGAACCTCCTGAC
 CTCGAGTGATCTGCCCCCTTGGCCTCCCAAAGTTCTGGGATTACAGACGTGAGCTACCG
 TGCCAGCTCAGTTTAGTAATGTATAACTGGGTTTTTACCAGTTGTAAATTACTCTTTTG
 TCGTGTTTTTTTGAGAACTGGCAATGACGGAGAACTAAAAGTGCCAGGCTGTTGCCTTG
 TTCCTGTTATTTTGCCTTAGTTTTTTTTTTTTTTTTTTTTTCTCTGAGACTGAGTCTTG
 TTGTGTTACCAGGCTAGAGTGGAGTGGCATGATCTCGGCTCACTGCAACCTCTGCCTCCT
 GGGTTCAAGTGATTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCGCCTGCCACC
 GCACCCGGTGAATTTTTGTATTTTGTAGTAGAGACGGGATTTTACCATGTTGGCCAGGCTG
 GCCTCGACCTCCTGACCTCATGATCCACCAGCTTCGGCCTCCCAAAGTGCTGGGATTACA
 GCGGAGAACCACCGTGCCCGTCTTGCTTAGTTATTTCTTGTTCCTCCTCTAGTCCTA
 TAGTTCTCTGACTGTATTGAGGAAATGTAATTAATATTTATTATGTTAATAGATATTTAT

FIGURE 14.10

29/64

GTGGTTGAATATTAGAAAATTCCTTATTTTGGTCACATATCCTGATCAGTAGTTGGTCTTC
TGGAGATAGTGATTTTTCTACTAGAGATGACTTTAGGACCTATTCAGGTTTTTTTTTAAGAT
CCCAATTTAAGGAAAGACTATTCTCATTATTGATTTTGCTATATGCAGGGAAAATTTATTT
CGAAAGGTTTTTCAGTTGGCTTTTTAGGGAAGATTATATATTCTCTTTTTTTTTTTTTTGGC
CTTTTCCCACATGTTCTAAAAATGATATATTCTTTAACTCCTATGAAAAATACATTGTTTC
AGTAATTGAAGATGCTGATTAAAGTCATATCTCTACACATTTTTTAAAAATTTGAGATAGA
TGGGACTTTGTCCCTTCTTACACCATTCACTTATTCCTTGGAAAACTATTATCCAATA
CTTATGTGGCAGACACTGTTTCTGGCACAAGGGATTGAGCAGTGAACAAAACCTGCCTTTT
TGGAGTTTACATTCTACTAGTGAAAGCGACAACAAGCAGATAGACACATTCAGTATATA
ATTCAGTGTGAGATGGTGGTGGTAAGTCTATGTAGGAAGAAAAGCAGGGTAAGGAGGCT
TGGAGTAACTGGAGTGAGTCATAGATGGACTTGTGAGGAAAGGGTTTCTGAAGAGGTGGT
ATTTGGGCAGAGATCTAAATAAAATGAAGCAACAAGCCATGAGAATATCCGGGGGAAAAT
GTTCTGGGCAGAAGCATCAAGCATAGAACTTGTGGTATGATATTTATTCTAGCACACATT
AATTTTAAAAATGTATAAAAGACATCCATTTAATCATATTAAAGATTTCCATGATTCATT
TAGACTTAGTCAGAAACCAAATTTATATTTTCTTTTTAAATAATTTTATCTCAACTCTTA
TTTTACCCAATAGGGGCCAGAGTTACTCAGCAAATACATTGGAGCAAGTGAACAAGCTGT
TCCGGATATTTTTATTAGGTTGGTAGCCTATGAATGTTTTTAAAGTAACTGACTCTGTTA
TTATTTATCAATCAGTGCTTTTTTTGGTCTTGTTTTTGAAGAACTGATATTTGAAACCT
GTGGTTTATGTGAATTATTAATAAGCTAGAGGACGTGGATTCTCTATTTTCATCAAATAAT
ACAAAACATTTTAGATATTAAATTTTGGAAATTATTTGGTTTTGTTTTACAATAGAAATA
CTCCTCAAAGTGGAATCGAAGTGTTTATTCAAAGAAATCTCAGAGTAGATTCTTATATGA
AGCAAATAATTGCCCTAATTTATCTCTAAATTTTGTAAAGTTCTAAATTTCTTTTTTCCCC
CAGTTTCTAATTTATCTCTTATAAGTCAAGAGTCCATCTGGCCAATTTAATTTTCAGTGAG
TGTAATTTTGCATATATTAAAAAACTGTATATGAATACAGAAGATGGTATTTAAGGA
TGAAATAATTATTCAAATGTGATAGCATTATGGGGAGTTTTTAAATAAAAGTTACTGTT
TTATTTCTCCAAAATTTTATTATAAAGTATACAGTTAAGAGAATATACATAAAATACAT
ATGCAGCTTAAGGAAGAATAATAAAATGAATACTTCATGTATTCACCACCGAGTTTACCA
GGAAAAAGCATAAACAATAAAACCTCTTCCACGTAATTCCTGGGTTAAAGAGAAGTTAT
AGTGGAAAATATTTGGGAGCAAACGATAATGAAAATACTATCCATTAAATTTGTTAGATG
TTGCAAACTGATTTCAAGGAAAATTTATAGTGTTAAATGTTTAGAAAAGAAAAAGGTT
AGAAGTTAACCCTTATGTATCTATCTCATGAAATTAGGAAAATTATAGATATAAACTAA
AAAATATGTTAAAGGGGAAATAATAAAGATAAGAATGAAGTTTAAATGAAACACAAAACAG
AGAAGCTCACAAAGCCAAGATTTATTTTTTGAACACCGAGTACAATTGACAAATCTCTAA
CAAGTTTGATTAAGAAAAAAGAAAGCATGAATAAAACAATTTTAGGGATAAAAAGGGAAAC
ATCGCTAAAGATATCCCAGAAATGTAAAAGATAATAAGGGAATATTATGAAAATATTCAT
GCCAATACATTTGAAAACCTTAGGTGACATAGACAAAAACAAAATTGACCAAAATTGAGCA
AAAAAGAAACAAAATCTGAGTAGTCCTGTAACTTAGTAAAAATTGAGTTAGAAAAGTTAA
AGAAGTCTTTACACAAATCAAACATCAGACTCAGTTTCTAGGAGAGTTTTGCCAAACAT
TCAAGTAGCAGATAATTCTGGTCTATTTTTGGCCCCAGAAGATATATTTTACTTGCCATG
CATTTAATGAGATAGCTGTTGATTTTTTTCAATCACCGTGACAGGTGTTTTATATTAGGT
GTTATTCGCCAGACATCTAGTCCACCTGTTGCCAGATATGGAATTAATATTCACTTATTT
TGAATTAATAATTTGTTAATAAATTAATAAAACAAAGTCAAAGTTCAAATTATTAAAAAAG
TAAAAGAAATAAATATATTTTTATAGAGAGCCCTTACAAAACAGTACCAACATAATGAGC
TTTCCAAATTTTGAATGGGCAAAATAAATGAATAGGCATTTACAAAAGAAGGAAGGGTG
GCCAATAAGTATATATTAATATAAAAAATGGTTACTTGTAATAGGAATCAAAGTGTTTGA
CTTATTGACTAAGAGTCAGTTTTTTGTTTTGATCCCTGTTAGTCTATCCAGAAGGCATGGG
TCTTAATAAACACCTTGACCTCAACAGTTTACTGAATACAAGGGTAATTTTCATATGCCTT
GCCTTCTTTAAGGGTTTGTGTAAAGATTAAAATAAATACATAAATATATATAAATACAT
TTATATGTATTTATATGTAATTACATACAACCTTGCTTCTTTAAGGGTTTGTGTAAAAA
TTAAAGAAGTATATAAATATATATAAATACATAAAATAAATACATTTCATATATGTATAT
GAAATCACTTTGCCAACTATGAAGCCTGATTCAAATATGAAATGTTGTTTGTTCCTCA
GAGCACAGGCTGCAAAGCCCTGCATTCTTTTCTTTGATGAATTTGAATCCATTGCTCCTC

FIGURE 14.11

30/64

GGCGGGGTCATGATAATACAGGAGTTACAGACCGAGTAGTTAACCAGTTGCTGACTCAGT
 TGGATGGAGTAGAAGGCTTACAGGTAATAATTATAAATACAGAAATAGAATGTTATAAC
 AAAATGTCATCATGTGCATCAGATTTTGGTAAAAAAATGTTCTTTTTTCTCTAGGTGTTT
 ATGTATTGGCTGCTACTAGTCGCCCTGACTTGATTGACCCCTGCCCTGCTTAGGCCTGGTC
 GACTAGATAAATGTGTATACTGTCTCTCTGATCAGGTGACAATTTTCATATTTAGAGT
 CCAAAACCCAACAAATGCTACACTCTTTCTTGTGAGCTTTACTTCTGCCAGGTAATGGC
 AATTGTCCTTAGAAGACCAGCTTTCTTAGGGAAAAAGCTTTAGCCACTGTTTGCTCAAAGC
 ATAAAAAGATTCTGAATTAGATGCAAAGCCTTTTTTTGGCCCAGTGCAAGTCTGAAAAC
 TTGTAATCCTTCTGTGTTGGCTGATTGGGGAAAAAAATGCAAGAAACCTAATGTATTA
 TATTTTCACATTATCTTCTGTTCAAAGATTACATACTTCCATTATCCTGTCAAAAAA
 ACTCTGATACAGAATCAAGCATGTGAATCGTAAGCATGTAAGCAGGTTTCATAGAGATAA
 TTTTTCAACTCTTCTTGTCTGTGTTGTTTCCAACCTCTTATTCTCCAATTTAGAAGCAA
 CAAATAAATGAATGAAAGAACAGATAGACAAATGAATAGTCAAAGGTATAAAGTATCTGT
 ATATATGTTACATGTAGCTATTATTTAAATTATTTAGATTTTCTTTTGAAAATACCTTCT
 TGGCACACTTGCCATAAATCTAGAAAAAAGCACTGTGTGAATAAGAAATTTTACACTG
 AATATTTTGTAGGTTTTTGGGTTTTTGTTTTTTCAGACAAGGTCTCACTTTGTCAACCCAGG
 CTGGAGTACACTGGTACGATCACAACCTCACTGCAGCCTCTATGGCCAGGCTCAAGCAAT
 CTCCCCACCTCAGCCTCCCGAGTAGCTGGGACCACAGGCACACGCTACCATGCCAGATA
 ATTTTATTATTAATTTTGTATAGAGATGGGGTCTCCCTGTGTTGCCAGGCTTTCTTGA
 ACTCCAGGGCTCAAGTGATCCTCCACCTCAACCTCCCAAAGTGTGGGATTACAGGCGT
 GAGCCACCATGCCCCAGCCTTAAGAGTGTGTTGATTTTCATTCATTTTCTATATATATTAT
 TTCTGTTGGGGAAAAAATTTCCAAGGAAGATAAATAGTAGGCTGTTGGTACATTTCTCAAC
 TTACTTATAAAGCTTTTTAGATATATAAGGTTAATTTATGAAGAAAAATCATAAGATACAC
 AATTTAAGATAATATTTTAAATTTTATTTTTTATTTGTTAAATAAAATTTTTCTCCTTTCA
 GGTGTCACGTCTTGAAATTTTAAATGTCCTCAGTGACTCTCTACCTCTGGCAGATGATGT
 TGACCTTCAGCATGTAGCATCAGTAACCTGACTCCTTTACTGGAGCTGATCTGAAAGCTTT
 ACTTTACAATGCCCAATTGGAGGCCTTACATGGAATGTGCTCTCGAGTGGACTCCAGGC
 AAGTTATATGAGGAAGTTGTTATGACATTTTATGAGTGATAAAAGAAGTACAATGTCAAA
 ATTTCCACCTTAAAAAATGCTATTTTTTAAACAACCTTTGGTAAAACCTGTATAGAAACATA
 AATTTACCTTTAGTTGAATGTTCCATAGTTGGAATATGGGTTTTGCAGAGAATTTATAAT
 TATGAAGTTTGATGTCTGTTTCTTTAACATTACCTTAATATTGGCAAAAACATGTTGGTG
 TTTGCAAGGATATTATTTAAATTTGGGATACCATGAATTAATACTACAAACAAAAATAAT
 TAGAGTTTTTTGTTTGTGTTTACTTTTAACTTTTAAAAAATAATCAGTTAAAGTTGTTGTT
 TTGAAGCTCACATTGTTCCAATCTGGCCAATAGGAGCCCTTTTGTATGGCTCCTGTATC
 TTTATGACATGTCCTCATCTTCTGAATCACTTCTCACTTCCAGATACAGTAAGTTAT
 TCTTGGCCAGGTGCAGTGGTTCACGCCTGTAATCCAGCACTTTGGCAGGCCAAGGCAGG
 AGGATCATTTGGGCCTAGTTTGAGACCAATCATGGTTGCACAACTGTACCCACTATGG
 ACAACAGAGTGGGATCTTGTCTCTGTGAAAAATTTAAAAATTAGCTGGGCATGGTGGCAC
 ATACCTGTAGTCCTAGCTTCTTGGGAGAGGCTGTGGCAGGAGGATCGCTTGAGTAAATCC
 AGGATGCAGTGAGCCATGCTTGTGCCACTGCACCTCCAGCATGGATGACAGAATGAGACCC
 TGCCCCAAAAAAGAAAAATATTCTTGGTTTATCTTGTACTTTCTGTATCCAGCCCTAG
 CATCAGCCTTTTCTCTAAAGACAGTATTATGATTTTAAATTTTACAGTAGATATTTGAAC
 TGTTACATTATAGACTTTACCATATATTTTTCTAGGAAGGATTATTCTATTACTCTTCTTT
 ACCACATTTGTTTGAATGTCTACAGAACCTACAGTTTCTAAATCAGAACTCCCTAGGT
 TTTTGCTATTTTGGCAAGCCATTGAAGTTCTTCCCTCTCCCTTTACTACCAGAAAGGTGT
 GTATTTGTAGAGCTCTCTATAATGAGAAAGCACTCTATAACATGGTTGATTATCATTTTT
 GGAGTAGAAAAGTATGAATGGAAAGTCAGAGACATAAAAAATAAGCCCAGAGGTCTGAGT
 CTTAGCTTCATTACAGACTTTCTTGGGGGATGGTTGGTAAATTATCTACACATTCTATCT
 TGTCTTTATAATTTTAAATAGTTAAATTTTACCATGTGCCCTCAAACCGTTAGAGAATTA
 ATGAGCTCTTTGAAAAATGCTTCTAAGTTTCTTGTATTGCTCTAATAGAATGCTATCTAT
 GTTATTATTTATTTCTGAGACTAAAAATTGTTTACATCTTTAAACTGGTTGTCTTTTGTG
 TATTTTAGGATGGAAGTTCCAGCTCTGATAGTGACCTAAGTCTGTCTTCAATGGTCTTTTC
 TTAACCATAGCAGTGGCTCTGACGATTACAGCTGGAGATGGAGAATGTGGCTTAGATCAGT
 CCCTTGTCTTTTAGAGATGTCCGAGATCCTTCCAGATGAATCAAAATTCATATGTACC

FIGURE 14.12

31/64

GGCTCTACTTTGGAAGCTCTTATGAATCAGAACTTGGAAATGGAACCTCTTCTGATTG
 TATCTTGTGCAGTCATCATTATACAGTTCTGAAATATAAAGCTATATGTTGGTGTAAGT
 TGCAGTGATTTCTCTCCTAACCCAGCCCCACATATTCTTCTGGTTGGTTGGTTCTTCAGT
 AAAATAGTCTTGTCTTCTGCTTACACTAATTGGTAATTTGCATTCTTGTAAAGATTTTC
 AAGACAGGGCTGGGAGCAAGGAACCAAAGTAGCGCGTGGTTGTGATTACCTTTGGTTTCT
 TTGAGGTTTCTCTTACCTAGTGGCTTTAAACATCTTTAGGAGCAGTTCCATTTTATAGT
 AAACCTAAATTCTGTTATCATGAACAGTTGAGGATAATGAATAATTTGATACAATAATGT
 AAGAAATTCCTGAAAACAAAGTGTTATCTGTGATACTTTTGTGTCATAGTAAGCACAAATG
 AAGTGTACTGATAATGTTTCAACAGGAAAGTGTTTGTATTAAATGTGGGCAGTATCACTG
 TTCTACTAGCATTCAACATCTCTTCTAAAAATTAATAGTGGTTCAGTGTAATTTTATTGG
 TACATGTAACATCTGTACATGTGTTTGGTTATCTATATGTTTCTGGTTTTTTGTACATT
 TGCTTTATTAATTTAGGCTTTTTTTTTTTTTTTTTTTTGTAGACAGTCTCACTCTATCATC
 CAGACTAGAGTGAGTGGCACAATTATGGCTCACTGCAGCCTTGACCTCCTGGGCTTAGG
 TGATTCTTCCACCTCAGCCTCCTGAGTAGCTGGGACTACAGGCACATGCCACCATGCCCA
 GCTAATTTTTGTATGTTTTGTAGAGACGAGGTTTACCATATTGCCCAGGCTGGTCTCAA
 ACTCCTGGGCTCAAGCTATCTGCGTGCCTTGACCTCCCAAAGTGCTAGGATTACAGGTGT
 GAGCCACTATGCCTAGCCTAACTCAGACTTTAAAAATATAAAGCAATTCAATTTTATTC
 CCAAGAACAGTAAGGTGGTGGTTTAAATTTTAGTCTTTAATTCTGTTTTTAATTTATTCTA
 TTTAGAAATGTCCCAGAACTTAGTATAACTTTACTTTCTGAAAATGAAGAAACCTGTCC
 TTGGGCATTAGTGTGTTGGATTTAAGCAACAAAGTTAAAAAACCTACCCTGTGTTATGG
 CAATTTTCACTTGATGGTGGTTCTATAACACAGGTATCAGTGAACCTTTATAAAGATGA
 ACAACTTTTTCAGCTTGCTTAATTTTCAGTTAATTAACATGTATACTTATCTATGTTAATGT
 TTTATTGCTTAAATGTTTAAATTTTATATTTTGGTAAACAGATAGTTTTTCTCTCCCCC
 TCTTCTTCCATCTTTTCACTACTACAATTTACCATGCAGAGCTCACAAATGTCTCTTGCA
CCAAGCTCCATGACTCAGGATTTGCCTGGAGTTCTGGGAAAGACCAGTTGTTTTACAG
CCTCCAGTGTTAAGGACAGCTTACAAGAGGGTTGCCAAGAACTTACACAAGAACAAAGA
GATCAACTGAGGGCAGATATCAGTATTATCAAAGGCAGATACCGGAGCCAAAGTGGAGTA
 TGGCTTTTTCCCCCTCATTATAATTGTTAAACTTCTTAAAAATTGTTTCACCCTTTTGA
 TATATATTTCTTTGACTTATAAACGAGCTATATTTATAAACAAGGGACCAGAACACATTA
 ACTCAGTCATGGTTATGTGCTTCTTGTCTTCAATGTTTCATTATCTTATAAGGAAGAGA
 ACGTATGGTCTCTTGAAAAAAGTGAACAATAAGAAGTAACAAGTGGACTACCACATTTTTT
 TTTACATCCTTAATTTAACTCTTTCGTCAATTTCTTTTTTTTACTTAAGGAGGACGAATCCA
TGAACCAACCAGGACCAATCAAAAACAGACTGGCTATTAGTCAGTCACATTTAATGACTG
CACCTGGTCCACACAAGACCATCCATTAGTGAAGATGACTGGAAGAATTTTGCTGAGCTGT
 AAGTAACAGATTCTGTTTTTGAAGTACAGCTACTATTACAAGTGACATAGTATTACACTT
 AAACCTTTAAAGTTCGTGTTTAAAAATAAAAAATATTTTGAATATTTAAAAGCTAATTCAA
 AAATATGTGTCGTAGCTATGCATTAAAAAACCCCAAAATGTCAGAAGTACAGAAGTCAAA
 ATTGAGTTTTTCAATTAACAGTTCATTGATTATATTTGAATTATTACATAATGGACTCATT
 TAATTTTAGTAACCTTTGGGCTGGGTGCTGCTGCTCATGCCTGTAATCCAGCTCTTTGGG
 AGGCCAAGGCAGGTGGATCACCTGAGGTGAGGTGAGGAGTTCGAGGCAAGCCTAACCAACACGGG
 GAAACCCCATCTCTACTAAAAATACAAAAATTAGCCAGGTGTGGTGGCATGTGCCTGTAG
 TCCCAGCTACTTGGGAGGCTGAGACAGGAGAATTGCTTGAACCCAGGAGGTGGAGGTTGC
 AGTGAGCCGAGATTGCACCACTGCACTCCATCCAGCCTGGGCCACAGAGCGAGACTGTGT
 CTCAAAAAATTTTTAGTAACTTCGAAGAAATAAGAAGGAAAATTTAAAGT
 TGAAAGTGATTCTAATGTATAGTTTATAAAATTTTGTATATAAAATACCTGTTTTGCCTT
 CAAAATAATTTATATTAATATTTTATTGACCTCAAGAACATTTAAATACATTGAGTTTA
 TTCATTTGTGGACCACATTTGTTATACATTGGATTTAAAGGATCCTTGCAATTGAGTTTA
 TGGCCACCTATGCATCTGAGACCCATGGACTGGGAACCATTTAGGTCAATGATTGAGTG
 TGATTCAATTTAAGAGATGTTTATTCCTGGTCTTTAGAAGCTGCTACCTTTTGTATCTA
 ATTTTGCAGTACTTTGAAGTATGTATGTATGTGTACATACGTTAGTGCTATGTATTTATT
 AAAGAAGAATCAGAAAACAGAGGTAAGGAAAAATAAGGAAACAAATTTCTGTTAAGCCCA
 CCACCTCCCAAAGCATATTTGTTTATATGCTTATATATGTTTTCTATTATGGTAAGAAC
 AGTCTGTACATATTGCTATATAGCAGTCCCCCTTTATCCACATACATCCTGAAAATTGTT
 TTACATTTTAAATGTTAACTACTTTATTGTTTTTAAATGTCATTTTATAGTGTAGCTATG

FIGURE 14.13

32/64

CCACAATATCCAATTTTTTAGACATTTAAATTGCTCCCAGGCAATGTGGTAATGAACATTCT
 TTGCAGCTGAATATATGCACATATCTAATTGTTTCACTAGGATAGAGGTGGAATTGTATA
 ACAGGGAGCTCACATTTTTTAAGGCTTTTGAAATGTATTGCCAAATTGCCTGCCAGATAT
 ACTGCACCATCCTAACATTGTGTGTTGCAGTATTTTTCTAACTTGGCCCTTTTGATTT
 TAGAAAAATGATATCAATAATTTACATTTCTTTGATTAAAGTGTAGAAGTTATAATTTTT
 CATATTATTCATTGTCTTTGTATTTTATCTTTTCTAACTTGTCTCTTCATCCCCCTTTGC
 TCCGTTTTCTATTGGAGTGAACCTTTATTTGTAAGAATTCTTTTTAATTTCTGTGACTGG
 AATTTTTTTTTCTAGTTTGTATTTCCTGTTTCTTAAATATAATTGTGTTTGCCA
 ACAATCCATTATCTTTTGTGTTTGTAAATGGTAGTATTTATACATATTAAATTATCTCTTTC
 TTTTTTCAGATATGAAAGCTTTCAAATCCAAAGAGGAGAAAAAATCAAAGTGGAACAAT
GTTTCGACCTGGACAGAAAGTAACCTTTAGCATATAAATATACTTCTTTTTTGATTTGGTTCT
GTTAAGTTTTTTTGATGGCTTTTCCATATGTTGTAACAGGAAAAAATGGTGTCTATGAAT
TTCTTCTTAATTTAACAAATTTGGTTAATTTATAAAATCACAGATTGGTAAATGCTATAA
TTATGTAATGATCAGGATTGAGATTAATACTGTAGTATAAATTGGGACATTATAACAGAT
TCCATATTTTATTTCCCTAAATCTAAATTCAGTCTTTAATGAAATAATATTAGCCAAATG
GTGGAACATAATTTATTTCTTTTGAGGAAAAGATAATAAAGAATGTAATTTAAATTTAAAT
TCTTGGAATCCCAGTTGTATATTCATCACCTTTGTAGCATTTGACAAATTTTATGCTTA
GCAGCTTCTTCACTGTTTTGAAATAAAATATCCTATTACCTACTGATACAATTATCTGTT
 CTTTGTATATCAAAAAATGTGAAATTTACACATAATTCAAATACATTAAATTATCCGCTC
 AACCAGAAATGAAATCACATCCCTCTACTATACATCCAGCTCCAAGCCCAAGATATT
 TAAATGACATCCATTCTCTCTCTTCCAGTTATGATTTTTATCTTTGATATTCTCTCATA
 TATGAACTAAATTTATAAAGTTAGCCACCATCAATACAATCTGCGTATCTAATATCTTAAC
 TATATAGTAATGGGGTAAGGGAACAGCAAAAGGAGAACATTAATTTAAATATACAAAGTA
 AGCCTGGGCAACATAGTGAGACCCCATCTCTTAAAAAATAATAGCCATGCATGATGGT
 ATGCTCTAGTCCCAGCTACTTGGGAGGCTGAGGTAGGAGGATCACTTGCTCCCAGGAGG
 TTCAAGGTTCTAAACCAGCAAAGCTCAGAATCCCAGGGGATAGAAACAAAGACTTAGTGG
 ATCACTAGTATTAACCTGAGACACGTCACCTGCATTGCACCTTGTCTCAGTTCTTTG
 ATGAAATCACTGAGCTGACATACCTGCCCTCTTTTACCATAAAGTGAGTTTCATGATCA
 GAAGCAATGTCTATGGGATAGCCTAACAAACAATGTAAAAACCATTTAGTAAGTTCATGA
 AGGGTGGTGGTGGTAAAAATTTGGAGAACATACAAAACAAATACAATTTCCAAGGTGTGTC
 CCTCCAGGAAGGACAAATTTGCTGCCCTGCTCTGTGATAGAAGAGGATCAGATGTAATCAA
 CCTGCCGTGCACTTGGGCTGTTCTCTCCTGGGTGTGGACTTGCTGGTTGGTCACTGCT
 GCTGACAAAGTAGGCTGTCAATATAGCTGGGTTGTCTGTCAGCTGTGGTGAGGGGGAAGT
 CCACATTGTGGAGGCCACATCCCTGCACTCTTGGCCAATTTGACCATGAATCTTAAGCAC
 TGGGGTGGCTGGAAAAGACAGCCGATTGACATCCATACAGAGGTCATCTTGACCACTTGA
 TTAGTATAAGCACTGAAGGCTTTTAACTGAGCATTACATAGGACACAAATATTCTGATT
 CTTTGGGCCCATTCCAAGAACTCTGGGCATACCTTTTCTCCAGACCTCATACCCAGTTGT
 GTTCTTTCCAAATTTCTGGTCATCTGGTTATGTTATTAGCCACTATCTGTGAATCAGCAT
 AGATTTTATATCAGACATCTCTACCTCCTGACAGAATGGAGGAGATATGTTACTTAACA
 ATTCTGTTCCCTTGGAAGATTTCTGTCTCCACTGTTTGTAAAGGGCTACTCCCTCAATGT
 AGCAGTAATGCTTTCACCTCTGATGGGAAGTCACAGTGGAATTTCTGGGTCTCCAAGATTA
 GTGTTAGTGATACACAGTGTCTGATAATCCCCAGAGTGTCTGGTGCCCTTGGATCCTGT
 GAAGAAGGCTTGAGAGAAAAGAAGATTCTAGGCAAGAAGTTGTGATGTGATGACAGGGCCT
 TTTCTCTGGCTCTTCATTCTTAGTCTGACCTAGGTGTGAGAATTAGGTGAGGGGCCATGA
 CTATATTGTGGTGACTCAAACCAGGCCCTTTGTTTACTAACTGGGAGATTTTACATTGTA
 AGAATCAAGTAGGATCTTTGCCCATGTATTTTGGTCTTAAGAACACAAATGATATGGCTC
 CAATGACTGGAGGAACACCAGGGTCCTTGGTCTCACGCTGATTTAGATAAAACGACTGTC
 AGGCCTCTGAGCCCAAGCTAAGCCATCTCCCTGTGACCTGCACGTATACATCCAGATG
 GCCTGAAGTAACCAAGAATCACAAAAGCAGTGAAAATGGCCTGTTCTGCTTAACTGA
 TGACATTTCCACCATTTGTGATTTGTTCTGCCCCATCTTAACCTGAGCGATTAACTTTGTGA
 AATTCCTTCTCCTGGCTCAAAACCTCCCCACTGAGCACCTTGTGACCCCCGCCCCCTGCC
 CCTAAGAGAAAACCCCTTTGATTATAATTTTCCACTACCCACCCAAATCTTATAAAATG
 GCCCCACCCCTATCTCCCTTCGCTGACTCCTTTTTTCGACTCAGCCCGCTGCACCCAGG
 TGAAATAAAAGCCTTGTGCTCACACAAAGCCTGTTTGGTGGACTCTCTTCACACGGAC

FIGURE 14.14

33/64

AAGCTTTAGTAGAGATCTCAAAAAATGGTTGGATGGTAGCAAATTACTAAGAACTCTCAAA
GTTTCTAAAGCCTTAGTTTCAGCTTGCTAGAAAAACCTATGTTGAGTATTATGGCTAGTTC
CATAGTTGAGTTGGGAAATGTCTTTGAGGAGACACTTTTTCACTTTGTATTTCATCTGTAC
ATTTTCTGTTACTTGCATTCTGTCTAGCTCAGGCTATTAGAGCAGGTACATTTTTATAAC
TGGAATGTTTATGTGTAGTGAAGCTCTGAGAGGACTTTGCATTAGATCTCAGCAGCATAA
TCAGAAGGTTGTCTTTGTCTCAGCAATTTTTAAGCTAATAGTAGCAGAAATTGCAGTGG
AAATAGACTGCTTTGCCACAACATTCAGAAAAATCATTTATCTTTTTATTGCAGTTCTTGT
CACCAAACAATACATTTTAGTACTTCTCAAATTGCAGAACTCTCATAGGGCTGGGAAAAAT
GCCTGTAGACACATACATACTATGAATGTGCTAATGTTTTTTGTATTTTCATAGCCCATC
AAAGCTCCTGAGTCAGTTTCCACTATAATCACTGCAGAATCAATCTTCTACAAGGTAAGC
TTTTGTAGAGTTACTGAAGGAAGAGTTGGGCCTAGTGGGTAATGTGCCACTAAAAATGTTG
GATTAGTCTAAAGGTCTCTGCTACTCTTTATTTGTATAAGGTGTGATTATACTTTTTGT
CCCTTCTTAGCTGTTTTCCCCCATAAGTGGCTGTTATTTAAAAACATCTCATCTAGAGCTGA
AGTGGGAGGAGAAAGTGCCTACTGACACATGATGTGAGGATCTTAAGTATTTTTTTTAG
TGTAAGATTGTAGGAATTATTCTTAAATGCTGATTGTATAGTGTGGAGCCATGGAAGACT
GAGCCGTTAGTGCATGGCATTGAAGAATGAGAAGGACAGAGACAGGATTGGACTAGTA
GAGGTTGTGCGACTGTGGTGTCAAATGGGTAGAGTAGGCCAGAGATTCTAAATGCCTTT
AAGTGGAGTTGAGCTGAGTAAGGGCAGTAGTGAGGATTAACACCTACTAGAAATTCATAG
TGAGAGGAATTCCAAGATGTTTTGATAAAAGAATGAGGAGGTCAGGTTTCCAGGGCCAA
AGTCCATGAACATCTGATACCTCAGTGAGAGAAGTGACAGATTGTTGTGTTTAAACCAGA
AGTCTTAGGAAAGGAATTAGAACATAGACCCCCAAGGCTCGGCAGGCCTGGCACGGCACA
GGCAGCAACCATTGAAGGCTATTTGGTGTTCGGGATCTGAACTGTCAATTTAGGGGACAG
TGGTGTGAGTTAGTACTTTATACTTGACCCAGGTGGACTGAGAACTCAAGTGATGATGC
CCTTAAGTATACTTTTTTTTAAAGCCACAATCTATATAGTCGAAGTCTGTTCTCCCAAC
AGGGGTACACTGGCATTCTCAGCAGGGCTGGGAAAAACCAACAACAAAAAAGTCTGTA
CACAGGCAAACATCTCTCTTATTTTTCCAACATTTAATACATTGTTAATAAAATATCTAA
AGTTTAGCAAACAGTTGCTGTGTATCAGTGGCTGAGCATTTTGCATGCTTTATTTTCATTC
AGTTCACTCTATGAGGTGGATACTACTATCCCCATTTTCTAGATGAGAACATTGAGGCAC
AGCGAGGTTAATTAACCTGTCCAAGATCACATAGCCAACAAGTCATGGAGTGAGGCAGTC
TCATGCCAGAGCTTAAGCCTAGAGCATAGTTTCTGGCTCTACAGCTTTAGCAAGTGACTG
GCTATGTGACGAGGACCAACCTCTCTAATGTCTCATCTGTAAAAATAGGAATTGTAAATAG
TTACTACCTCAGTGGGTCAAATGAAATCATATGTGTTAAGCACTTAGCAGAGTAAGCACT
CAATGAATAGTAGGAGTTATCACATCTTCGTATTTGTGCATTACCTTCACAGTTTACAGA
TTAAGGCCAGAAGCAACTTGTGTGAGCTACGGGTTTAGTGTAACAGTTTCCATGTGTG
TCTCCATGGAAGGGTGTGTGGGACCTGTTATTGTGACTGTCTGTACTTTTGTATTGTGT
CTGCCACCCATGTTTATTAAATGATAAGGACAATAATGCAACAAAGTAGTCAAGTAATGT
TGCAAATGCCCAGTATTGTAGTGGCTATCACAGCAGTGCCACTGGCAGGCAGCACCATGG
TGGCAAGTTCAAGAGGTCAGTCCAGCCACTGAGCTAGAGCCAGATCAGGCATGCAAGA
GGAGCCTGAGTGGGAGCCACTGGGGATCACGGCCAAGAGTGTGACCACCAAGACCCAGA
ATGGCTGAGTGGCCTCCCTGGAGCATGGCAGTGGCAGAACTCCATGAACTCAGATCT
GGTGATGCCTAAACTAGTGCTGTTCTCGTGTGGACCCCTTTTCTCTACCAGAAACCTTGA
ATCCTCTCAGCAAATGAGGAGACTACTCAGATCAGTGACTTAGTCCTGTTTGGTGTATA
TATGTGTACACAACACAGCACATATTAATAAATACCTACTATGTGCCAGGCACTGCCTAC
CACTGGAATCTTTCACCTAAGACATTGTTTTTACTTTGCATTTCTGCCTTTACACTATGAA
AGTAGATGTTTTGGATTATATTCATTTCAGCATAACATTTGAATATGCTGTGTATGCATA
GTAAGCCTATGATAAGCAAGTATTCTCATTTAGAATTTGGGAATATTGATTATACATGTG
GACAAACAAACCATAAATGCAAACTATTTATATGATAAATAACTTTGGACTGATGGCTGG
GAGGAAGGACCAGCTATTGATGGGTAGGAAGTAGCAAGTAGCGGACTGTGGCCTGCATAG
ACCAGACCCATCCGTAGTGATCCAGATGAAACAGCCACCCCTCAGACACTTGGATAAAGGG
TCCACCAGGAAAAAACTCCTGGCCTATCAGGTGCTATGTTACAGTTCAGTTACTGGAAGT
ATTTCTCAAAAGTGTTTTTATGGTTGAGGTACACATTCCTACAGCTTTACCTGCTGCCA

FIGURE 15.1

34/64

AGTCCCTGTTTCAAGGGAAGCAGCAATGAATTACACTGTTCCCGTAGTCAAGGACAGTAT
 ATCTTACCAAGAACTATACCCACTTAAGGAGGTGCTGGATGTCATAAAGATTTGGATCAA
 CCATTATGGGTGTTTCAAGGAGAGATTATTTCCAGCTCAAGACCCAGGGAAGAGGACATA
 GGATGGATAACCAGAGTCATAGGGAGGATTTAACACAGGACATGTACACATTAGTTAGTTG
 GGTATAAAGTGGAACAGAAATGAATGAGACACAAAGCCTTGAATGCCAGAAATACTAGTA
 GTCCTGTTGTGGAAGGATATAAACTCAACTGGGAGTGGAAGAGAAAGGCAGCAGTGAGT
 CTAGGAGATGTACAGTAGGTTGAGGTAAACATATCCTGAAGACTATAATCCAAAGATTAT
 TTTTGGTTTGAATTTGTTTTGGTTTGAATTCATGGTATCTATTTTCTTTGAGTGGATGGT
 TGGGGAGGGTGGCATGTAGAATGCATTCTTACCAAATCAGCATGATTTTCAAGACAGTAC
 AGAGAAAAGACTGCTGAGCTGATGTAGGAGCTTTGGCTGCAGTCTCTATGGCTTTCAGCA
 AGCCGTTTAACTTACTACTGCTTCATGACTGTGGCTAACAAAGTAGGGATAGTACGGAG
 CACAGAGGATTTTATAGGCGGTGAAACTATTAATACTCTCTTTGTATGATACTATAATGG
 TGGGTACATGTCAATTATACATTTGCCCAACCCACAGAATACACAGCACCAAGAGTGAAC
 CCTAATGTGAACCTCTGGTCTTTGATGATGCTATGTCAGTGTACGTTTATCCGTGTAACAA
 GTGTACCACTCTAGTGGTGGGAGGGGTATTGATAATAGGGGAGGATGTGCATGTGTGGG
 GGCAGGAAGTATATGGGAAATCTCTCTACTTCTGCTCAATTTTGTCTGTAAACCTAAAACC
 TCTGTAAAAAATAAAGTCTATTTTAAAAAGTGGGGATGGTATTACGGCAATATAAAAT
 CAAAACTCTTTTGAACAAATCTTTTCTCCAGATGTAACTGTCTATATATGCACCCTCGT
 ATGTGTATGTATAATTTTCAATTCAAACGTGAAACAACCTTTAGAATTGGCACCAACATAT
 AAACACTGATACATTAGACTATCTCGAACACCTTTTACTGACCACCTTTGAAAACCTTGCTT
 ACCTATTAAGGTTTCAATTCATAGCTGTGATGTTCTATTTTATTTTCAATGTGGGATTATC
TTCTGTTTCCCCCAGGGAGTATATTACCAAATTGGTGATGTTGTTTCTGTGATTGATGAA
CAAGATGGAAAGCCCTACTATGCTCAAATCAGAGGTTTATCCAGGACCAGTATTGCGAG
AAGAGTGCAGCACTGACGTGGCTCATTCCTACCCTCTCTAGCCCCAGAGACCAATTTGAT
CCCGCCTCCTATATCATAGGTAAGTTTGACAAATGGCACAGGTTTTTTTTTAACTTAGTT
 AACTCTCCAATATTATGTAAAAGAGTGTGTTAGTCAGCTTGGGCTGTGAGGACAAAATAT
 CACAGACTGAGTGGCTTAAACAACAGAAAGTCACTTCTCACAGTTGTGGAGGCTGAAGT
 CCAACATCAAGGTGCTGGCAACACGGATTTCTGGGGAGGCTTTTCTTCTTGGCATATAGA
 TGGTCACCTTCTTGCTGTGTCTCACATGGCCTTTTATGGAGTGAGAGCTCTTTGGTGTA
 TCTTCTTATAAGGACACCATTTCTGTGATGAGGGCCCCACCCTTATGGTTTCAATTTAA
 CCTTAATTGCCCTCCCTAAAGGTCTCATCTCCAAGTACCATCACATTGGGGATTAGGGCTT
 CAACATATAAATTTGGAGGGTGGCGGGGGGGGATGCAATTCAGTCCATAACAAAAAAGC
 ATGAGTATTATTAAGTACAAAAAATTAGAGAGCTTTATAGAAAATATGAGGCATTTTAT
 GTAGCTGGAGTGTGAGTGCATCTAGTTATTTTGTAGTTAGAGCAATGTGCATCTACTAAGA
 AGTGGTATGGATAAGATTTTTTGGAGTGACCCAGGGTTAACTGTACTACAAGAATGTA
 TTGCTCAGGAACTAGGTTATTTAGGTTACTTATTTATACAAACCTATTCAAAAATAATTT
 AGGAAAGAACTATCCAGTTATCCCATACTTGCAAATTTCTCAATATGTGTGCCCTCTGCAT
 GCTACACATGTCATCTTAGGCCTTTATAGTATAAAGGCTGATAGTTGAAATGGCAGCTGC
 TGTGCTTTTGTAAATTTCAAAGCTGCCAAAACAGTTGTGAGATAGACTCACAAGAATTTA
 CTGATTAATACAATTTTTAAAGTTTTCAGATTTTTTACAGTTACTTCAGACTTTTTATCTT
 TCTGCAGTGAGCATGCATCATTAATTTTGCATCCTGAGAACAAGCATAAGTGTGTTTTTG
 GAGAGAACTCCAGGGACAAATAATATACCACGTGTTATTTCTCACCTATATGTCAAGTTTGA
 TACATTACCAAAACAATTTCTAGCCTTCTGCTTATAAGTATATAGAATTTTTATTTACCTTA
 TCTATGGATCAGGATCTCAGCAGAGGCAGTGATGTATCAGAATCACCTTCGGGATTCCTC
 TACTGCCCTCTCTTTCTAATCCCCAGATTCTGATATGCATCCTTGTCTACAGCGAGGCA
 GCATGGCATGAGGTGAGAACACCAGTTCTGGAGCCAGACTGTCTAGGTTTACAGCCTGCC
 ATTTACCGGCCATGTGACTTTGGCAAGTTTCTTAGTCTCTCTTGCCTCACTTTCTCCTATA
 TGTAATAATGGGAATAATAAGTGCCTACCTCAGAAGGTTGATGTGAGGAATGAAGGTAT
 TGATACATGTAACTTAGAGCAGTGTGGGTACAAAATAAACATGATGCAAGTGTTCAATC
 ACTGTTTTTGGGAGAATGCCATATCTTTAAGCCGTAAAGAAGAAAAAATGATTAAGAA
 TAATTTCAAAGTAATGCATGTTTCAAGGGCTAATGCCAGGTTGCTCCCAGAGTGGTCTCT
 CCCAGTGTCTAGAAATTTAACATCTTATGAAAATGATATATATGGTCAAAAATGTATTT

FIGURE 15.2

35/64

AACCTTTCCCTTGGCTGCCTTCCAGGGCCAGAGGAAGATCTTCCAAGGAAGATGGAATAC
 TTGGAATTTGTTTGTTCATGCACCTTCTGAGTATTTCAAGTCACGGTCATCACCATTTCCC
 ACAGTTCCACCAGACCAGAGAAGGGCTACATATGGACTCATGTTGGGGCTACTCCTGCA
 ATAACAATTAAGGAATCAGTTGCCAACCATTTGTAGTTTACAAATTAATACTGGGTTTCC
 AGGCCTGGTGTGGTGGCTCACGCCGTGTAGCCCCAGCTATTGCACCACTGCTCTCCAAGCT
 GGGCAATGGAGTCAGATTCTCTTTCTTAAAAAACCAAAAAAACTGGATTTCCAGTTCT
 CTAATATTCTTAGTACCACAAGATATGTCATAGGTATCTTTAAATGAAATTCCTAGCTGG
 AAAAGTGACTAAAAAGTTTTCTCTGCTACCTAGTAATAAACAAATCATTGTTTATTAC
 TGGTCACTTAGAAAAATTAAGGGGATAGGGCCAGGCACAGTGGCTTATGCCTGTAATTGC
 AGCACTTTTAGAGGCCGAGGCAGGCGGATCACCTGAGGTCGGGAAGTGATCGCCTGAGG
 TCAGGAGTTTCGAGACCAGCCTGGCCAACATGGCGAAACCCCGTCGCTACTAAAAATACAA
 AAATTAGCCAGGTGTGGTGGCATGTGCCTGTAATCCCAGCTATTTGGGAGGCTGAGGCAG
 GAGAATCGCCTAAACCCAGGAGGTGGAGGTTGTAGTGAGCCAAGATTGCACCCGCTGTGCT
 CCAGCCTGGGCAACAGAGTGAGACTCTTGTCTCGGAAAAAAAAAAAAAAAAAAGGCTG
 GGCACAGTGGCTCACGCCCTTAATCCCAGCACTTTGGGAGGCTGAGGCAGATGGATCGCC
 TGAGGTTGGGAGTTCGAGACCAGCCTGGCCAGCATGGTGAAACCCTGTCTCTACTAAAAA
 TACAAAAATTAGCAGGTGTGGTGGCGCACACCTGTAGTCCCAGCTACTCGGGAGGCTGA
 GGCAGGAGAATTGGTTGAACCCAGGAGGCGGAGGTTGCAGTGAGCAGAGATCGTGCCACT
 GCACCTCAGCCTGGGTGGACAGAGCAAGACTCCGTCTCAAAGAAACAAACAAAAAATTA
 AAGGGATAGAATATAATGAAATATATTTTGAACCTTAAATTATATTCTATATGTGTATCTT
 CCTAGGCAAAAGCTGTAATTTCCAGAGAGACCATTAGGAACAGGTAGTATCTATTTTTCT
 CCATTATTTATTTCTAGAAACTCATAAAATGGATTGTATTTTTCTATAAGAACAAAAATAT
 TAATTAAGGTATAGATGACTGACCAAGGGCTTAATCAAATAAAATGACTAACAGCATCTA
 TCATAAAGCCACACAAGCCTTATGTTCTCATCTCAAAAATGCTGTGACAGCTTTTTGGCT
 GCTTTAACCATAAGAAAAATGATTGGTGGATGATTTTATTAGCCAGGCTTTTAAAAACT
 TTCATCTAGGCCACGTGCGGTGGCTCATGCCTGTAATCCCGGCACTTTGGGAGGCTGAG
 TGGATGGATCACTTGAGGTCAGGAGTTCAGGACCAGCCTGGCCAACATGATGAAACCTG
 TCTCTACTAAATATACAAAAATTAGTTGGGTGTTATGGTGCATGCCTGTAATCCAGCTA
 CTCGGGAGGCTGAGGCAGGAGAATTGCTTGAACCTCGGGAGGTGGAGATTGCAGTAAGCCG
 AGATCGTGCCACTGCACTCCAGCCTGGGTGATAGAGCAAGACTGTCTCAAAAAAGAAAAA
 AAAGAAAAAATTTTAATTTAATCCTTCTGTAGAAACAGGCATTTCAGAACCATTCATTGA
 TCTTAATAAAGCTGCTCTTTACTGTTTCTAGTCAAAAATGAGACTTCGATCAAAACATAA
 GATTTTATACTGCAGATAGTCAGCTTCACCAAAGCCGAGAGGAAACATGTGAGATCAG
 GCTTCTGCTTGATAGTCTCTTGACTACCATTAAACGAATATTGGGAGGTTCATGAAAGT
 CATTGGTAGGCCATTAGCATTGATATCTTTAAACATCTACCCTAAACCATCTGCTATGG
 ACCCATAATAAGAGGCCTGTTGTATATGAAATTGTCTAGAATTCAGGTGCAGGTCTTGC
 CGGTTAAGTAAGGGAGCAACACGTAAAAATGGGAGAGGAGTGGGGTGTACTCACTTGCCTC
 CTCTTTTGTCTGATTTAACCAGCATTTTTCAACCTTGGGAAAAATTTGCAGAATCTAAGT
 TGATTGTAATGATTTTGAGCTGCAGCAGCTTTAACTCTTACCCTTTTTCCACATAGTTAT
 GGTGTTTGAGTTGGAAAAGAAACAACTATAGGTAGCTACACGTACATAATTATCTCTTTAT
 TCACAAAGGGTATAGTAAAAATTGATTGTAAATAACTTTCTAAGTGCCAATATTCAAAACT
 TTTGGATTAAAAATGTATTTTTACCGTGCATTTACTTTGGATGTATTTATTTTCAATTTAA
 CAATTTAAATGGGGCTCTTTAACCAAAAATGGTATTTAAAAACAAAAACAGTATCGTACTT
 AGAATTTGGAGTAGAGGCCGGGCACAGTGGCTCACGCCGTGTAATCCAGCACTTTGGAAG
 GCTGAGGCAGGCGGATCACCTGAGGTGAGGAGTTCGAGACCAGCCTGGTCAACATGAAAC
 CCCGTCTCTACTAAAAATACAAAAATTAGCTGGGCGTGGTGGCGTGCGCCTATAATCCCA
 GCTAGTCTACTCGGGAGGCTGAGGCAGGAGAATCGCTGGAACCTCAGGAGGCAGAGACTGC
 AGTGAGCCGAGATCGCGCCACTGCACTCCAGTCTGGGTGACGGCATGACTCCATCTCCAA
 AAAAAAAAAAAAAAGATTTTGGAGTAGATTTCATCATTAATAAGTAACAGATTTTAGGAAA
 ATCAAAAAATGGCTAATAAAAAATGAACACAATGTAAAAACATTTATTAAAAATGTAGACTTTT
 AAAAAATCTATAAATGATCATCTGTTTATAAATTTGGCAGATGGTTGTGTACCATCTTTTA
 AAATAAAGATTGAATTTACCCAGTGTGATGGTTCCCATTTGCTTATATTTCTCTGCTGA

FIGURE 15.3

36/64

GGCCGGACCTGATATGGCCCTGGTCTGTGTTCCACGCTTGTTTCCTCATTACCACTAAA
 ATCTTTCCCTGTATGCCCCGCCAATTTTTCTGGCTCTGAGTCCTTGTTCACTGTCTCT
 CTCCAATTCTACCTTCCAAAGGCCTTTCTTAACACCTTCGGATTCTTTCTTTGAGAACTT
 TCCAGATTCCCATGCCTTTTGGGAATCAATCTCTATCTATTGTCATCACATTTAAGTTT
 CTACTTCCATCATCCTCACTCCTATCCCTTTGGTCCTGGGATGACAGGGATGCTGTGTTT
 TATTTACTCATCTTTGTAACCTCCACATAACCTAACCCCGGTTCTTGCTTATGGGAGATG
 CTGATTGTAGGGTCTGAGTTAGATACTGTAACTAAAATGCTTGTTGATATTTTAGTTAT
 TAATTCATATTAACCTTTGGCTGAACTTTTAAATTCTATTGTGAATAGTCAAGTAAATTT
 TAGATTGTTACATTCTGGGTTAGTATTAGATTGTTTTTAAGATTGTTTTAAACAAGATGT
 TTTTAAGATGAGTTTTAAATAGTTCTCTTAACACAAAATAAAGCTTAATATGAGTATTTGA
 AGGAAATTATCCCAAACCATTCCAGTTCCCTGGCTGTGAAAGGCTTTTCCAGGCCTAATAA
 GTTTTCCACTTCAGCCGTAAGTAGGTGAAATCAAATGAACAATAGAGGGAAATGTATTTA
 TTTGCTTTATACACATGCATGTGTGTTGTGTCTACATATAAACATTGCACACGCTTAGAA
 TGAAGTTTCTGTCTATGCCCAGAAAAGGGAGAGGCATTTTGTGGATTTTGTCTGGCTGCC
 CTGGGGATGTTTGAAGAACTGTGTCTGTTTACTTACATACCAGGTGTGTGAGCCATACCTTT
 GGTAGGAGGTATACCTCCTACACCCAAGAAATATAAGCCAGGAGAAGGTCTGTGCCAAG
 AGAAGGAACCCAAATGACCCACAAGAGGTGGGCCATTAATTATTGGGTGAGATGCATAAA
 TGCACAGTAATTTATTTAAGCACCTCTTAATGGTGACCCACAAGGAAGATTGCTCGTAGT
 AGCGGAAAGGTTTCAATAAAATAAGAGAAAAAAGCAGAATGTAGAATGTATGATAGCAA
 TTCTGCAAAACAAGAAGCATCTTTTATAAAAGATGGAAGGAGCCCAGGCACAGTAGCTCAT
 GCCTGTAATCCCAGCACTTTAAGAGGCTGAGGTGGAGGATCACTTGAGCTGCAGTGACCC
 ATGATTGTGCCACCACTCCAGCCTGGGTGATAGAAGTGAGACCTTCTCTCAAAAAAAAAA
 AAAAAAAAAAAAAAGACGGAAATTCCTCCAGAATTTTAACATGTCAACAGAGGTTTCTGC
 AGCTACTTTTTTTCAGCTTTTATACTTCGCAGTATTTTCCAAATTTTCTCTAACAAGCAGTA
 TTTTCCAAATTTTTTACAATAAGCACACACACACACACGTTTGTTTGATAAGTGCCC
 AACTGGTGGTGAACAACCGCTGGCTTTTAGTCTATACATATCTAGAATATTTTATAAATA
 GTAGTTCTTAAACCCCTTGAAAGGGAGTGAATGACCAGCTGAGAAAATAAAGTCAGTGATT
 TCATTATTTTTCTATATTCACATCATGATTCTAGGAAAGAACTTGGGAGTGACTTCCTTC
 AGCTTCAGCCACTCCTGGGCCAGGCGCATGCTTAGCTCTGTGGTAAAGGTCAACAGCTTC
 TTCTGCAGGGTGCCTGTATCATCTGAATTGGAGGTTTGGCGAGGGTAAGAGACTGATGTA
 GGTTCAAGTTTTTCTTTCTGTCTCCACTTGAAATCTGTCTTCCCTTCCAGACTGCCTG
 CGCTGCTGACTTAAGGCCCCAACACCAAACACAGAAGCAACAGCCTTACACAGAGTGTTT
 AGCAAGCTCCAACAATTGTGTAAGGTAAAGTTTCTTTATAGATTCTTTTCTATATCGC
 TCCTAGTGTTCTGTTTCTCTGATCGAATTCGCTGATAACAGTTGCTGAGACTCTGAA
 AGAAGGCAAGGAAGTACTGTTTCTCATTATAAACTGTTTGAATTTTGGCCATCTT
 TTTGCTATGAATATGTAGTGCTTTGATACATTTTTTAAATCAAAAAGTAATGAAAGAGAT
 CACATAGGGAAAGATAGATTGGATTATTTTTTAAAGTTTATATACTAAATTGAAAAGCAAA
 GAATAAAATGGGAGAAACAGCTCCCTCATGTGGCTGTTGGCAGGAAGCTTCCATTCTCT
 CTGTGGGCCTCCACAGGTTTGCTCACAGCAAATGGTCCGTGACAGAAAGACGCAAGGGCA
 GTTGACCCCAAGATGGAAGCCACCATCTTTTCTATAACCTAATCTGAAAGAAGGGACATA
 CCAGCACTTCTGCCATATGCTGTTGGGTACACAGACCAACTCTGGTACAGTGTGAACAC
 AGGACCACACAAGGGCGTGAATTCGAAGGGCAGAGACCACTAGGGACCACCTCAGAGGCA
 CAGAGGGACACCCTATCCAGCTGGTGGCCAATGTAAATTAACATAGCTTTTTTAGAATAGC
 AATATGTATCTATAATCTTAAAGTATTTAAAGTACTTCTTGATCCAGTAATTTTCAATTC
 TAAGAATCCATGCTAAGAGGATTTAAATGTGGACCAAAAAATGGGTATAAAAAGAAGTT
 GTTAACAGTATTTAAAGTTGTGAAAAACCAGAAACAATCTAAAGGTCCAACAATAGGAAA
 ATGAATTTTGATATTTTTCTAATAGAATTTTATGCTGTCTATCAGAAATACCATTTACAAA
 TAATTTTTAATAACGCAAAAAAAGTTTATAAAATGTTTAGTGTAACCTGGACACAAC
 TACATAATGATTCTGATTTTGTAAAAAACAACAAAAACACACATATACACATGCA
 TACATATGCATATAAAGAAAACTGGAACAAACAAAATAACAAGCATAGTTGGAATTACAG
 TCATTTTAATATTCTTTATGCTTTTAAAAATTTTGAAGTTTGTATTACTAGCATCCACTA
 CTTACGTAGTCAGGAAAAAATAACAACCTTTAAATAGATATTTAGGTCCAAGATGGTAA

FIGURE 15.4

37/64

TCTAAATGGTGTGTACAGGCTGAATGTGTGCCTGATCCCCATGCCCCAAGTTTCATATGTTA
 AAGCCCTGGCCCCCAAGGCAATGGTATTAGGGGAGTAGGGCCTTTGGGAGGTAATCAGAT
 TTCTACGAGGTCATGAGGGTGGAGCCCGCATAGTGGAATTAGTGTCTTTTAGGAAGAGG
 AGAACAGACCAAAGCCTTCCTTTCTCTCCTCACTATGTAAGAAGACAGCCAGAAGGTGGC
 CACAGCCAGGAAGAGAGCTCTCACCAGAACCCAAATCTGCTAGCACCTTGCTCTTGGGT
 CTCAGCATCCAGAACTGTGAGAAATGAATGTGTGTTGTTTAAACCACTCAGGCTACGGTA
 TTTTGTGTGCAGCAGCCCAAGCTGACAGAGATAGAAACAACACAAGGACCCATCAGCAGAC
 GAATGGATGATCAAAACGTGGTGAGGTCGTGCAGTGGGATATTATTAGCCGTAGAAGGA
 ATGAAATTCGTATACATGCTATAATGATGAACCTTGAAAACATGTTAATGGAAATAAGCC
 AAACCTTAAAAGGACAAATATTGTATAATTCCACTTATATGAGTTAGTTACCTAGAATAGG
 CAAATTATGTCATAGATACAGAACATTAGAGGTTACCAGGGTTGTGGGAAGAGGGGTATT
 GTGGGTACAAATTTTCGGTTTGGAGTGATTTTGAAAAAATTCTGGAAATGGGTAGTGACA
 GTAGTCAACATGATGAATGTACTTAATGACACTAAATTTGTACACTTAAAAATGGTTAATA
 CTGGGCTGGCGCAGTGGCTCATGGCTGTAAATCCCAGAACTTTGGGAGGCCAAGACAGGC
 GGATCATGAGGTCAGGAGATTGAGACCATTCTGGCTAACATGGTGAAACCCCTGTCTCTAC
 TAAAAAATAAAAACAAATAAAAAAATAATAGCCGGGCATGGTGGCAGGCACCTGTAGTC
 CCAGCTACTCGGGAGGCTGAGGCAGGAGAATGGTGTGACCTGGGAGTCGGAGCTTGCAGT
 GAGCTGAGATCGCGCCACTGCACCTCCAGCCTGGGCAACAGAGCCAGATTCCGTCTCAAAA
 AAAAAAAGAGTTGATACCTGGGTGCGGTGGCTCATGCCGTGAATTTTTCAGCACTTT
 GGGAGGCCAAGGCAGGCAGATCAGTTGAGGTCAAGAGTTAAGGACCAGCCTGGCCAACGT
 GGCGAAACCCCATCTCTATTAAAAATACAAAAATTAGTCGAGTGTGGTGGTGGGTGCCTG
 TAGTCCCAGCTGCTGGGAGGATGAGGCCTAGGAATTGCTTGAACCCAGGAGGCAGAGGTT
 GCAGTGAGTTGAGATTGCGCCACTGCACCTCCAGCCTGGGGACAGAGCGAGACTTAGTCT
 CAAAAAAGAGTTAAAAATTGTAAGTTTGTATTATGCATATTTTACCATAATCTTTAAAAA
 TAGATATATAGGAGATAAAGTCAACAGAATTTAATAACCAGTTGTAAATAGAGACTGAGT
 GAGGAGGATGAATTAAGGAAGACATTGAGTACAACCTTTTGGTAGGTGAAAAACTCTTAA
 AAAAAACGTGGGCAAAGATCCTACTTGATTCTTATAATTTAAAAATCTCCAGTTAGTA
 AACAAGGCTAGGTGGAGATTTGCATGTGATGTGAGGTGTGTGTTCTGTTTTGTAATGTGA
 GGACTGTGAGCCATCTCCTGGACTTGAATATCCATTAGATAATTGAAAATACGGATTGTA
 GAACTCAGGAGACGTGCAATGCAGTAACAAAACCTCTGCACCTAGTTGATTTCTGTCTCCT
 AATTTAATGCTTTTATGGGACAACTGTTAGGCAGGTGGGCAAGATGGACAGCCATATTT
 TTGTGGGTTTCTGGCCTGTGGGCCAGCCTCAGTGCTCACTCTGAGGTCATGTCCAACTT
 AGAACACATTACAGGCCTACCACAGTCAAGGCTCCCTTTCTCAACTCTAGTCCTCTGCACA
 AATATCCGAAGCCTAGAAATAATAATCATCTGTCTTGTGTCTTGCAATTATGAAAGCCTA
 GGAAAGGGCCTTGGGAATTAAGAAGATGGAAAACTGGTCTAACTGCTGCATGCTTCAG
 CTTGCAGGGGAATCACTGAAATGGGGACAGGCCATAAAAGGACAACCCAGAAGAGTGGCTT
 CAGCAAAGGCATCGTTTTTCAGAGCAAGCTAGAGAATCCTGCCAGCGTCTCAGGCAGGG
 CCCCTGGGCACAGAGGTAGGCAAGGGAGTGTCCAGCATGTTGATGCCCTGAGCATCAG
 AATAATGCCATAGAGGAGCTTCCAAAGAGTTTCAATTCAGGTTTTGTAAAGCCGAACATTC
 TAGGCAAAATAAAATTTGATTTTGTGAATAAAGCTTGTCTTCTCAACTCCAGTGCAGATT
 TCATAGATTGATAGTGGCTTGTGATCCAGATAAAGAAAAACAATTTTCAAAGATTTCATAT
 TCTTTGTAGATGTACGGATTTAGAGACCATCTAATCTAACTCCCTCATTCTACAGATAGG
 AAAAAATGAGGCCTAAAGAAAGTTAAGAAAAATACCATGGAAATGTCACTGCTGAACTGCCAT
 ACGTAGGATCCGAAAGAAATTTGGGTAAATGCTACTGTGAGAAATACAGTACTAGGTCCAA
 AGAATCTAATACAAATTAATAATCTAAATGTTATTTCTAAAGCATCCCTGCACATGGCTG
 AACTTACATAGTTTCATTTTCTTTCTTTTCTGTTGAAGAAGAGGCAATTGGCTGGGTGCA
 GTGGCTCATGCCTGTAATCCTGGCACTTTGAGAGGCCGAGGCGGGTGGATCACCTGAGGT
 CAGGAGTTTGAGACCAGCCTGGCCAACATGGTGAAACCCCATCTCTACTAAAAATACAAA
 AATTAGCTGGCTGTGGTGGCCGCTGCCTGTAATCCCAGCTACTCCAGAGGCTGAGGCAGG
 AGAATTACTTGAATCTGGGAGGTGGAGGTTGCAGTGAGCCAAGATCAGGCCATTGCACTC
 TAGCCTGGATGACAAGAGGGGAACTCCATCTCAAAAAAAAAAAGAAAAAGCAATCACT
 AACCTGTGTTGTTTATTAAACATGACAGACTGGCATGAAGTAATTACCAAACTGTAAACA

FIGURE 15.5

38/64

AAAAAGCTACAATCTGCCAGGCATGGTGGCTCATGCCTGTAATCCCCACCTTGGGAGGC
CAGGTTGGGGGATCACCTGAGGCCTGGAGTTCAAGACTAGCCTGGTCAACATGGTGAAAC
CTCGTCTCTACTAAAAATACAAAAATTAGCCCGGCGTGGTGGCACATCCCTGTAATCCCA
GTTACTCAGGAGGCTGAGGCAGGAGAATCACTTGAACCTGGGCAGTGGGGAGGTTGCAGT
GAGCCAAGATCGCACCGTTGTACTCCAGTCTGGGCCGACAGAGTGAGACTCGGTCTCAAA
AAAAAGAAAAAGAAAAAGCTACAACCTTAATCTCAACTTCTCATAACATCATCTCTACTT
CTGATTAGAAGAGTGGAAGTGGGGAGGTTTATTACAAAAAGACTGTTATACCTTACACAC
TTCTCCCCATGAATAGTGAAGGTGTGAGTGAAAAAGACAGCAATTTTATTTTTTTTTTGA
AACAGGTTCTTGCACTGTCAACCGGGCTGGAGTGCACTGTTGTGATCACTGCTCACTGCA
GCCTCCACCTCCCAGGCTCAAGTGATCCTCCTACCTCAGCCTCCTGAGTAGCTGGGACCA
CAGTTGTGCACTACCATGCCCAGCTATTTTTTTTTTAAGAGATGGGGTCTCACTATATTGC
TTAGGCTAGTTCTCAAACCTCCTGGCCTCAAGCAGTCCTCCGACCTTGGCCTCCCAAAGGG
TTGTGATTACAGGCATAAGCCACCACACCCAGCCAGCAGTTTTAGAATAAAGGGTGAAGG
TGCTGTTGGGGAAATATAATTTAAAAAACAAAATCTTCTCTCAACCCAGAAATCCTCTCC
ATGAAGGCAGTAGAGAAAGATAAGCTTTATTATTGAATAAAAAATTAAATGAGAATGTGAT
GCACATCACAGGCACCTTTGCTAAGAGATCACAAAGACAGAAGGAAATTCACCATTTTGT
ACAGCCAAGCAGGTACAGCCCATTACATGTATGTTTTCGAGATAAATAGTCCTCAACTAA
GAGAACTTGACAGCACCCTGGTCACACAGTTCATTCTAACTTTACCTGATAATTGATGT
GACCACTTGTGTTATCTAAGATATCAACTTTTCGGGGGTGGGGGAGTGTTGGAAACAGGAG
TTACTTTTTATAGCTTGGTGCAAGGTACTCATTAAGATTAGGCTGTTACCCTCCACAGAA
ACTGGAAGATAGGTATGCTATCTGGTAATGTTTACATTTCCCAGATCCTTGAGAAAGACA
TTCTTAGGTCATAAAGCTGACAAAAGGCTGATTTCAGTTTTTAAATATATATATCTGTATA
TGTATTTCA

FIGURE 15.6

39/64

```

actgagagacaggactagctggatttcctaggctgactaagaatccctaagcctagctgg
|||||
actgagagacaggactagctggatttcccaggccgactaagaattcctaagcctagctgg

g-aagggtgaccacatccacctttaaacacggggcttgcaacttagctcacacctgaccaa
| |||||
ggaagggtgaccacaccctcctttaaacacagagcttgtaactcagctcacacccgaccaa

tcag-----agagctcactaaaatgctaattaggc-aaagacaggaggtaaagaaa
|||
tcaggtagtaaagagagctcactaaaataccaattaggctaaaaacaggaggtaaagaaa

tagccaa-tcatctattgcctgagagcacagcaggagggacaatgatcgggatataaacc
|| |||
taatcaaatcatctatcgctgagagcacagggggagggacaatgatcgggatataaacc

caagtcttcgagccggcaacggcaacccccctttgggtccctccctttgtatgggagctc
|| | |||
caggcatttgagccagatcaggtaacctcctttgggtccctcacactgtatgggagctc

tgttttcatgctatttcactctattaaatcttgcaactgcac--tcttctggtccatggt
|||
tggt-----ttcactctattaaatcttgcaactgcacactcttctggtccatggt

tcttacggcttgagctgagctttcgctcgccatccaccactgctggttgccgccaccgca
| || |||
tggtccggctcaagctgagcttttgctcgccgtccaccactgctgaatgccgccattgca

gacccgccgctgactcccatccctctggatcatgcagggtgtccgctgtgctcctgatcc
||| |||
gacctgcccttgacttccacccctccggatccggcagagtgtccgctgcactcctgatcc

agcgaggcacccattgccgctcccaatcgggctaaaggcttgccattggtcctgcatggc
|||||
agcgaggcacccattgccactcccgatcaggctaaaggcttgccattggtcctgcacagc

taagtgcctgggttcacctaattgagctgaacactagtcaactgggttccatgggttctct
|||||
taagtgcctgggttcacctaattcaggctgaacactgggtcgctgggttccacgggttctct

tctgtgaccacagcttctaataagagctataaacactcaccgcatggcccaagggttcatt
|| |||
tccatgactcacagcttctaataagagctataaacactcaccacatggcccaagggttcatt

cctt-gaatccataaggccaagaaccccaggctcagagaacacgaggcttgccaccatctt
| || |||
cgttggaatccatgaggccaagaaccccaggctcagagaataaaaggcccgcc-ccatctt

gggag
|||
gggag

```

FIGURE 16

40/64

TCCTGTGAAC CTCTAGAGGA TTTGCGCCTG CTCTTCAAAC AACAACCAGG AGGAAAGTAA 7860
 CTAAAATCAT AAATCCCCAT GGCCCTCCCT TATCATATTT TTCTCTTTAC TGTTCCTTTA 7920
 CCCTCTTTCA CTCTCACTGC ACCCCCTCCA TGCCGCTGTA TGACCAGTAG CTCCCCTTAC 7980
 CAAGAGTTTC TATGGAGAAT GCAGCGTCCC GGAAATATTG ATGCCCCATC GTATAGGAGT 8040
 CTTTCTAAGG GAACCCCCAC CTTCACTGCC CACACCCATA TGCCCCGCAA CTGCTATCAC 8100
 TCTGCCACTC TTTGCATGCA TGCAAATACT CATTATTGGA CAGGAAAAAT GATTAATCCT 8160
 AGTTGTCCTG GAGGACTTGG AGTCACTGTC TGTTGGACTT ACTTCACCCA AACTGGTATG 8220
 TCTGATGGGG GTGGAGTTCA AGATCAGGCA AGAGAAAAAC ATGTAAAAGA AGTAATCTCC 8280
 CAACTCACCC GGGTACATGG CACCTCTAGC CCCTACAAAG GACTAGATCT CTCAAAATA 8340
 CATGAAACCC TCCGTACCCA TACTCGCCTG GTAAGCCTAT TTAATACCAC CCTCACTGGG 8400
 CTCCATGAGG TCTCGGCCCA AAACCCTACT AACTGTTGGA TATGCCTCCC CCTGAACTTC 8460
 AGGCCATATG TTTCAATCCC TGTACCTGAA CAATGGAACA ACTTCAGCAC AGAAATAAAC 8520
 ACCACTTCCG TTTTAGTAGG ACCTCTTGTT TCCAATCTGG AAATAACCCA TACCTCAAAC 8580
 CTCACCTGTG TAAAATTTAG CAATACTACA TACACAACCA ACTCCCAATG CATCAGGTGG 8640
 GTAACCTCTC CCACACAAAT AGTCTGCCTA CCCTCAGGAA TATTTTTTGT CTGTGGTACC 8700
 TCAGCCTATC GTTGTGTTGAA TGGCTCTTCA GAATCTATGT GCTTCCTCTC ATTCTTAGTG 8760
 CCCCCTATGA CCATCTACAC TGAACAAGAT TTATACAGTT ATGTCATATC TAAGCCCCGC 8820
 AACAAAAGAG TACCCATTCT TCCTTTTGTT ATAGGAGCAG GAGTGCTAGG TGCACTAGGT 8880
 ACTGGCATTG GCGGTATCAC AACCTCTACT CAGTTCTACT ACAAACATC TCAAGAACTA 8940
 AATGGGGACA TGGAACGGGT CGCCGACTCC CTGGTCACCT TGCAAGATCA ACTTAACTCC 9000
 CTAGCAGCAG TAGTCCTTCA AAATCGAAGA GCTTTAGACT TGCTAACCGC TGAAAGAGGG 9060
 GGAACCTGTT TATTTTTAGG GGAAGAATGC TGTTATTATG TTAATCAATC CGGAATCGTC 9120
 ACTGAGAAAG TTAAAGAAAT TCGAGATCGA ATACAACGTA GAGCAGAGGA GCTTCGAAAC 9180
 ACTGGACCCCT GGGGCCTCCT CAGCCAATGG ATGCCCTGGA TTCTCCCCTT CTTAGGACCT 9240
 CTAGCAGCTA TAATATTGCT ACTCCTCTTT GGACCCTGTA TCTTTAACCT CTTGTTAAC 9300
 TTTGTCTCTT CCAGAATCGA AGCTGTAAAA CTACAAATGG AGCCCAAGAT GCAGTCCAAG 9360
 ACTAAGATCT ACCGCAGACC CCTGGACCGG CCTGCTAGCC CACGATCTGA TGTTAATGAC 9420
 ATCAAAGGCA CCCCTCCTGA GGAAATCTCA GCTGCACAAC CTCTACTACG CCCCATTCA 9480
 GCAGGAAGCA GTTAGAGCGG TCTCGGCCAA CCTCCCCAAC AGCACTTAGG TTTTCCTGTT 9540

FIGURE 17

41/64

AAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGACTGGCAACTGATTTTACCCACAAGCCCAA
LysLeuLeuGlnGluAsnLysGluGlnAlaIleThrLeuGluLysThrGlyAsn...PheTyrProGlnAlaGln
SerSerPheArgArgThrLysAsnArgProLeuProTrpArgArgLeuAlaThrAspPheThrHisLysProLys
AlaProSerGlyGluGlnArgThrGlyHisTyrProGlyGluAspTrpGlnLeuIleLeuProThrSerProAsn

ACCTCAGGGATTTCAGTATCTACTAGTCTGGGTAGATACTTTACGGGTTGGGCAGAGGCCTTCCCCTGTAGGAC
ThrSerGlyIleSerValSerThrSerLeuGlyArgTyrPheHisGlyLeuGlyArgGlyLeuProLeu...Asp
ProGlnGlyPheGlnTyrLeuLeuValTrpValAspThrPheThrGlyTrpAlaGluAlaPheProCysArgThr
LeuArgAspPheSerIleTyr...SerGly...IleLeuSerArgValGlyGlnArgProSerProValGlyGln

AGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATTCCCAGATTTCGGACTTCCCCGAGGCTTACA
ArgLysGlyProArgGlyAsnLysGlyThrSerSer...AsnAsnSerGlnIleArgThrSerProArgLeuThr
GluLysAlaGlnGluValIleLysAlaLeuValHisGluIleIleProArgPheGlyLeuProArgGlyLeuGln
LysArgProLysArg.....ArgHis...PheMETLys...PheProAspSerAspPheProGluAlaTyrArg

GAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTT
Glu...Gln...ProCysPheProGlyHisSerAsnProGlySerIleProGlyValArgTyrThrIleSerLeu
SerAspAsnSerProAlaPheGlnAlaThrValThrGlnGlyValSerGlnAlaLeuGlyIleArgTyrHisLeu
ValThrIleAlaLeuLeuSerArgProGln...ProArgGluTyrProArgArg...ValTyrAspIleThrTyr

ACACTGCGCCTGAAGGCCACAGTCCTCAGGGAAGGTCGAGAAAATGAATGAAACACTCAAAGGACATCTAAAAAA
ThrLeuArgLeuLysAlaThrValLeuArgGluGlyArgGluAsnGlu...AsnThrGlnArgThrSerLysLys
HisCysAla...ArgProGlnSerSerGlyLysValGluLysMETAsnGluThrLeuLysGlyHisLeuLysLys
ThrAlaProGluGlyHisSerProGlnGlyArgSerArgLys...METLysHisSerLysAspIle...LysSer

GCAAACCCAGGAAACCCACCTCACATGGCCTGCTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCA
385 395 405 415 425 435 445
AlaAsnProGlyAsnProProHisMETAlaCysSerValAlaTyrSerLeuLysLysAsnLeuGlnLeuSerPro
GlnThrGlnGluThrHisLeuThrTrpProAlaLeuLeuProIleAlaLeuLysArgIleCysAsnPheProGln
LysProArgLysProThrSerHisGlyLeuLeuCysCysLeu...Pro...LysGluSerAlaThrPheProLys

AAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAG
LysSerArgThr...ProIleArgAsnAlaValTrpLysAlaLeuHisAsnGln...ProCysAla...ProLys
LysAlaGlyLeuSerProTyrGluMETLeuTyrGlyArgProPheIleThrAsnAspLeuValLeuAspProArg
LysGlnAspLeuAlaHisThrLysCysCysMETGluGlyProSer...ProMETThrLeuCysLeuThrGlnAsp

ACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAACATTACAAGGAACCTAT
ThrAlaAsnLeuValAlaAspIleThrSerLeuAlaLysTyrGlnGlnValLeuLysThrLeuGlnGlyThrTyr
GlnProThr...LeuGlnThrSerProPro...ProAsnIleAsnLysPheLeuLysHisTyrLysGluProIle
SerGlnLeuSerCysArgHisHisLeuLeuSerGlnIleSerThrSerSer...AsnIleThrArgAsnLeuSer

CCCTGAGAAGAGGGAAAGAACTATTCCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCA
Pro...GluGluGlyLysGluLeuPheHisProCysAspMETValLeuValLysSerLeuProSerAsnSerPro
ProGluLysArgGluLysAsnTyrSerThrLeuValThrTrpTyr...SerSerProPheProLeuIleProHis
LeuArgArgGlyLysArgThrIleProProLeu...HisGlyIleSerGlnValProSerLeu...PheProIle

TCCCTAGATACATCCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTG
SerLeuAspThrSerTrpGluGlyProTyrProValIleLeuSerThrProThrAlaValLysValAlaGlyVal
Pro...IleHisProGlyLysAspProThrGlnSerPheTyrLeuProGlnLeuArgLeuLysTrpLeuGluTrp
ProArgTyrIleLeuGlyArgThrLeuProSerHisPheIleTyrProAsnCysGly...SerGlyTrpSerGly

FIGURE 18.1

42/64

GAGTCTTGGATACATCACACTTGAGTCAAATCCTGGATACTGCCAAAGGAACCTGAAAATCCAGGAGACAACGCT
 GluSerTrpIleHisHisThr...ValLysSerTrpIleLeuProLysGluProGluAsnProGlyAspAsnAla
 SerLeuGlyTyrIleThrLeuGluSerAsnProGlyTyrCysGlnArgAsnLeuLysIleGlnGluThrThrLeu
 ValLeuAspThrSerHisLeuSerGlnIleLeuAspThrAlaLysGlyThr...LysSerArgArgGlnArg...

AGCTATTCTGTGAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAATAAAATCA
 SerTyrSerCysGluProLeuGluAspLeuArgLeuLeuPheLysGlnGlnProGlyGlyLys...LeuLysSer
 AlaIleProValAsnLeu...ArgIleCysAlaCysSerSerAsnAsnAsnGlnGluGluSerAsn...AsnHis
 LeuPheLeu...ThrSerArgGlyPheAlaProAlaLeuGlnThrThrThrArgArgLysValThrLysIleIle

TAAATCCCCATGGCCCTCCCTTATCATATTTTTCTCTTTACTGTTCTTTTACCCTCTTTCACTCTCACTGCACCC
 ...IleProMETAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThrAlaPro
 LysSerProTrpProSerLeuIleIlePhePheSerLeuLeuPhePheTyrProLeuSerLeuSerLeuHisPro
 AsnProHisGlyProProLeuSerTyrPheSerLeuTyrCysSerPheThrLeuPheHisSerHisCysThrPro

CCTCCATGCCGCTGTATGACCAGTAGCTCCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATT
ProCysArgCysMETThrSerSerSerProTyrGlnGluPheLeuTrpArgMETGlnArgProGlyAsnIle
 LeuHisAlaAlaVal...ProValAlaProLeuThrLysSerPheTyrGlyGluCysSerValProGluIleLeu
 SerMETProLeuTyrAspGln...LeuProLeuProArgValSerMETGluAsnAlaAlaSerArgLysTyr...

GATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCACACCCATATGCCCCGCAACTGC
AspAlaProSerTyrArgSerLeuSerLysGlyThrProThrPheThrAlaHisThrHisMETProArgAsnCys
 METProHisArgIleGlyValPheLeuArgGluProProProSerLeuProThrProIleCysProAlaThrAla
 CysProIleVal...GluSerPhe...GlyAsnProHisLeuHisCysProHisProTyrAlaProGlnLeuLeu

TATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCTCT
TyrHisSerAlaThrLeuCysMETHisAlaAsnThrHisTyrTrpThrGlyLysMETIleAsnProSerCysPro
 IleThrLeuProLeuPheAlaCysMETGlnIleLeuIleIleGlyGlnGluLys...LeuIleLeuValValLeu
 SerLeuCysHisSerLeuHisAlaCysLysTyrSerLeuLeuAspArgLysAsnAsp...Ser...LeuSerTrp

GGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAACTGGTATGTCTGATGGGGGTGGAGTTCAAGAT
GlyGlyLeuGlyValThrValCysTrpThrTyrPheThrGlnThrGlyMETSerAspGlyGlyGlyValGlnAsp
 GluAspLeuGluSerLeuSerValGlyLeuThrSerProLysLeuValCysLeuMETGlyValGluPheLysIle
 ArgThrTrpSerHisCysLeuLeuAspLeuLeuHisProAsnTrpTyrVal...TrpGlyTrpSerSerArgSer

CAGGCAAGAGAAAAACATGTAAAGAAGTAATCTCCCACTCACCCGGGTACATGGCACCTCTAGCCCCCTACAAA
GlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHisGlyThrSerSerProTyrLys
 ArgGlnGluLysAsnMET...LysLys...SerProAsnSerProGlyTyrMETAlaProLeuAlaProThrLys
 GlyLysArgLysThrCysLysArgSerAsnLeuProThrHisProGlyThrTrpHisLeu...ProLeuGlnArg

GGACTAGATCTCTCAAACTACATGAAACCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTC
GlyLeuAspLeuSerLysLeuHisGluThrLeuArgThrHisThrArgLeuValSerLeuPheAsnThrThrLeu
 Asp...IleSerGlnAsnTyrMETLysProSerValProIleLeuAlaTrp...AlaTyrLeuIleProProSer
 ThrArgSerLeuLysThrThr...AsnProProTyrProTyrSerProGlyLysProIle...TyrHisProHis

ACTGGGCTCCATGAGGTCTCGGCCCAAACCCTACTAAGTGTGGATATGCCTCCCCCTGAACCTCAGGCCATAT
ThrGlyLeuHisGluValSerAlaGlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyr
 LeuGlySerMETArgSerArgProLysThrLeuLeuThrValGlyTyrAlaSerPro...ThrSerGlyHisMET
 TrpAlaPro...GlyLeuGlyProLysProTyr...LeuLeuAspMETProProProGluLeuGlnAlaIleCys

GTTTCATCCCTGTACCTGAACAATGGAACAACCTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCT
ValSerIleProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuValGlyPro
 PheGlnSerLeuTyrLeuAsnAsnGlyThrThrSerAlaGlnLys...ThrProLeuProPhe.....AspLeu
 PheAsnProCysThr...ThrMETGluGlnLeuGlnHisArgAsnLysHisHisPheArgPheSerArgThrSer

FIGURE 18.2

43/64

CTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACC
LeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPheSerAsnThrThrTyrThrThr
 LeuPheProIleTrpLys...ProIleProGlnThrSerProVal...AsnLeuAlaIleLeuHisThrGlnPro
 CysPheGlnSerGlyAsnAsnProTyrLeuLysProHisLeuCysLysIle...GlnTyrTyrIleHisAsnGln

AACTCCCAATGCATCAGGTGGGTAACTCCTCCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTTGTCTGT
AsnSerGlnCysIleArgTrpValThrProProThrGlnIleValCysLeuProSerGlyIlePhePheValCys
 ThrProAsnAlaSerGlyGly...LeuLeuProHisLys...SerAlaTyrProGlnGluTyrPheLeuSerVal
 LeuProMETHisGlnValGlyAsnSerSerHisThrAsnSerLeuProThrLeuArgAsnIlePheCysLeuTrp

GGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATG
GlyThrSerAlaTyrArgCysLeuAsnGlySerSerGluSerMETCysPheLeuSerPheLeuValProProMET
 ValProGlnProIleValVal...METAlaLeuGlnAsnLeuCysAlaSerSerHisSer...CysProLeu...
 TyrLeuSerLeuSerLeuPheGluTrpLeuPheArgIleTyrValLeuProLeuIleLeuSerAlaProTyrAsp

ACCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCT
ThrIleTyrThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValProIleLeuPro
 ProSerThrLeuAsnLysIleTyrThrValMETSerTyrLeuSerProAlaThrLysGluTyrProPhePheLeu
 HisLeuHis...ThrArgPheIleGlnLeuCysHisIle...AlaProGlnGlnLysSerThrHisSerSerPhe

TTTGTTATAGGAGCAGGAGTGTCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTAC
PheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIleThrThrSerThrGlnPheTyr
 LeuLeu...GluGlnGluCys...ValHis...ValLeuAlaLeuAlaValSerGlnProLeuLeuSerSerThr
 CysTyrArgSerArgSerAlaArgCysThrArgTyrTrpHisTrpArgTyrHisAsnLeuTyrSerValLeuLeu

TACAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTGCGCGACTCCCTGGTCACTTGCAAGATCAACTT
TyrLysLeuSerGlnGluLeuAsnGlyAspMETGluArgValAlaAspSerLeuValThrLeuGlnAspGlnLeu
 ThrAsnTyrLeuLysAsn...METGlyThrTrpAsnGlySerProThrProTrpSerProCysLysIleAsnLeu
 GlnThrIleSerArgThrLysTrpGlyHisGlyThrGlyArgArgLeuProGlyHisLeuAlaArgSerThr...

AACTCCCTAGCAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGT
AsnSerLeuAlaAlaValValLeuGlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCys
 ThrPro...GlnGln...SerPheLysIleGluGluLeu...ThrCys...ProLeuLysGluGlyGluProVal
 LeuProSerSerSerSerProSerLysSerLysSerPheArgLeuAlaAsnArg...LysArgGlyAsnLeuPhe

TTATTTTTAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAAGTTAAAGAAATTCGA
LeuPheLeuGlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGluIleArg
 TyrPhe...GlyLysAsnAlaValIleMETLeuIleAsnProGluSerSerLeuArgLysLeuLysLysPheGlu
 IlePheArgGlyArgMETLeuLeuLeuCys...SerIleArgAsnArgHis...GluSer...ArgAsnSerArg

GATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGG
AspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeuLeuSerGlnTrpMETProTrp
 IleGluTyrAsnValGluGlnArgSerPheGluThrLeuAspProGlyAlaSerSerAlaAsnGlyCysProGly
 SerAsnThrThr...SerArgGlyAlaSerLysHisTrpThrLeuGlyProProGlnProMETAspAlaLeuAsp

ATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTT
IleLeuProPheLeuGlyProLeuAlaAlaIleIleLeuLeuLeuLeuPheGlyProCysIlePheAsnLeuLeu
 PheSerProSer...AspLeu...GlnLeu...TyrCysTyrSerSerLeuAspProValSerLeuThrSerLeu
 SerProLeuLeuArgThrSerSerSerTyrAsnIleAlaThrProLeuTrpThrLeuTyrLeu...ProProCys

FIGURE 18.3

44/64

GTTAAC TTTGTCTCTTCCAGAATCGAAGCTGTAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATC
ValAsnPheValSerSerArgIleGluAlaValLysLeuGlnMETGluProLysMETGlnSerLysThrLysIle
LeuThrLeuSerLeuProGluSerLysLeu...AsnTyrLysTrpSerProArgCysSerProArgLeuArgSer
...LeuCysLeuPheGlnAsnArgSerCysLysThrThrAsnGlyAlaGlnAspAlaValGlnAsp...AspLeu

TACCGCAGACCCCTGGACCGGCCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAA
TyrArgArgProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProProGluGlu
ThrAlaAspProTrpThrGlyLeuLeuAlaHisAspLeuMETLeuMETThrSerLysAlaProLeuLeuArgLys
ProGlnThrProGlyProAlaCys...ProThrIle...Cys.....HisGlnArgHisProSer...GlyAsn

ATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCA
IleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer...SerGlyArgArgProThrSerPro
SerGlnLeuHisAsnLeuTyrTyrAlaProIleGlnGlnGluAlaValArgAlaValValGlyGlnProProGln
LeuSerCysThrThrSerThrThrProGlnPheSerArgLysGlnLeuGluArgSerSerAlaAsnLeuProAsn

ACAGCACTTAGGTTTTCTGTTGAGATGGGGG
ThrAlaLeuArgPheSerCys...AspGlyGly
GlnHisLeuGlyPheProValGluMETGly
SerThr...ValPheLeuLeuArgTrpGly

FIGURE 18.4

45/64

LysLeuLeuGlnGluAsnLysGluGlnAlaIleThrLeuGluLysThrGlyAsn...PheTyrProGlnAlaGln
ThrSerGlyIleSerValSerThrSerLeuGlyArgTyrPheHisGlyLeuGlyArgGlyLeuProLeu...Asp
ArgLysGlyProArgGlyAsnLysGlyThrSerSer...AsnAsnSerGlnIleArgThrSerProArgLeuThr
Glu...Gln...ProCysPheProGlyHisSerAsnProGlySerIleProGlyValArgTyrThrIleSerLeu
ThrLeuArgLeuLysAlaThrValLeuArgGluGlyArgGluAsnGlu...AsnThrGlnArgThrSerLysLys

AlaAsnProGlyAsnProProHisMETAlaCysSerValAlaTyrSerLeuLysLysAsnLeuGlnLeuSerPro
LysSerArgThr...ProIleArgAsnAlaValTrpLysAlaLeuHisAsnGln...ProCysAla...ProLys
ThrAlaAsnLeuValAlaAspIleThrSerLeuAlaLysTyrGlnGlnValLeuLysThrLeuGlnGlyThrTyr
Pro...GluGluGlyLysGluLeuPheHisProCysAspMETValLeuValLysSerLeuProSerAsnSerPro
SerLeuAspThrSerTrpGluGlyProTyrProValIleLeuSerThrProThrAlaValLysValAlaGlyVal
GluSerTrpIleHisHisThr...ValLysSerTrpIleLeuProLysGluProGluAsnProGlyAspAsnAla
SerTyrSerCysGluProLeuGluAspLeuArgLeuLeuPheLysGlnGlnProGlyGlyLys...LeuLysSer
...IleProMETAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThrAlaPro
ProProCysArgCysMETThrSerSerSerProTyrGlnGluPheLeuTrpArgMETGlnArgProGlyAsnIle
AspAlaProSerTyrArgSerLeuSerLysGlyThrProThrPheThrAlaHisThrHisMETProArgAsnCys
TyrHisSerAlaThrLeuCysMETHisAlaAsnThrHisTyrTrpThrGlyLysMETIleAsnProSerCysPro
GlyGlyLeuGlyValThrValCysTrpThrTyrPheThrGlnThrGlyMETSerAspGlyGlyGlyValGlnAsp
GlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHisGlyThrSerSerProTyrLys
GlyLeuAspLeuSerLysLeuHisGluThrLeuArgThrHisThrArgLeuValSerLeuPheAsnThrThrLeu
ThrGlyLeuHisGluValSerAlaGlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyr
ValSerIleProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuValGlyPro
LeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPheSerAsnThrThrTyrThrThr
AsnSerGlnCysIleArgTrpValThrProProThrGlnIleValCysLeuProSerGlyIlePhePheValCys
GlyThrSerAlaTyrArgCysLeuAsnGlySerSerGluSerMETCysPheLeuSerPheLeuValProProMET
ThrIleTyrThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValProIleLeuPro
PheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIleThrThrSerThrGlnPheTyr
TyrLysLeuSerGlnGluLeuAsnGlyAspMETGluArgValAlaAspSerLeuValThrLeuGlnAspGlnLeu

FIGURE 19.1

46/64

AsnSerLeuAlaAlaValValLeuGlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCys
LeuPheLeuGlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGluIleArg
AspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeuLeuSerGlnTrpMETProTrp
IleLeuProPheLeuGlyProLeuAlaAlaIleIleLeuLeuLeuLeuPheGlyProCysIlePheAsnLeuLeu
ValAsnPheValSerSerArgIleGluAlaValLysLeuGlnMETGluProLysMETGlnSerLysThrLysIle
TyrArgArgProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProProGluGlu
IleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer...SerGlyArgArgProThrSerPro
ThrAlaLeuArgPheSerCys...AspGlyGly

FIGURE 19.2

47/64

SerSerPheArgArgThrLysAsnArgProLeuProTrpArgArgLeuAlaThrAspPheThrHisLysProLys
ProGlnGlyPheGlnTyrLeuLeuValTrpValAspThrPheThrGlyTrpAlaGluAlaPheProCysArgThr
GluLysAlaGlnGluValIleLysAlaLeuValHisGluIleIleProArgPheGlyLeuProArgGlyLeuGln
SerAspAsnSerProAlaPheGlnAlaThrValThrGlnGlyValSerGlnAlaLeuGlyIleArgTyrHisLeu
HisCysAla...ArgProGlnSerSerGlyLysValGluLysMETAsnGluThrLeuLysGlyHisLeuLysLys
GlnThrGlnGluThrHisLeuThrTrpProAlaLeuLeuProIleAlaLeuLysArgIleCysAsnPheProGln
LysAlaGlyLeuSerProTyrGluMETLeuTyrGlyArgProPheIleThrAsnAspLeuValLeuAspProArg
GlnProThr...LeuGlnThrSerProPro...ProAsnIleAsnLysPheLeuLysHisTyrLysGluProIle
ProGluLysArgGluLysAsnTyrSerThrLeuValThrTrpTyr...SerSerProPheProLeuIleProHis
Pro...IleHisProGlyLysAspProThrGlnSerPheTyrLeuProGlnLeuArgLeuLysTrpLeuGluTrp
SerLeuGlyTyrIleThrLeuGluSerAsnProGlyTyrCysGlnArgAsnLeuLysIleGlnGluThrThrLeu
AlaIleProValAsnLeu...ArgIleCysAlaCysSerSerAsnAsnAsnGlnGluGluSerAsn...AsnHis
LysSerProTrpProSerLeuIleIlePhePheSerLeuLeuPhePheTyrProLeuSerLeuSerLeuHisPro
LeuHisAlaAlaVal...ProValAlaProLeuThrLysSerPheTyrGlyGluCysSerValProGluIleLeu
METProHisArgIleGlyValPheLeuArgGluProProProSerLeuProThrProIleCysProAlaThrAla
IleThrLeuProLeuPheAlaCysMETGlnIleLeuIleIleGlyGlnGluLys...LeuIleLeuValValLeu
GluAspLeuGluSerLeuSerValGlyLeuThrSerProLysLeuValCysLeuMETGlyValGluPheLysIle
ArgGlnGluLysAsnMET...LysLys...SerProAsnSerProGlyTyrMETAlaProLeuAlaProThrLys
Asp...IleSerGlnAsnTyrMETLysProSerValProIleLeuAlaTrp...AlaTyrLeuIleProProSer
LeuGlySerMETArgSerArgProLysThrLeuLeuThrValGlyTyrAlaSerPro...ThrSerGlyHisMET
PheGlnSerLeuTyrLeuAsnAsnGlyThrThrSerAlaGlnLys...ThrProLeuProPhe.....AspLeu
LeuPheProIleTrpLys...ProIleProGlnThrSerProVal...AsnLeuAlaIleLeuHisThrGlnPro
ThrProAsnAlaSerGlyGly...LeuLeuProHisLys...SerAlaTyrProGlnGluTyrPheLeuSerVal
ValProGlnProIleValVal...METAlaLeuGlnAsnLeuCysAlaSerSerHisSer...CysProLeu...
ProSerThrLeuAsnLysIleTyrThrValMETSerTyrLeuSerProAlaThrLysGluTyrProphePheLeu
LeuLeu...GluGlnGluCys...ValHis...ValLeuAlaLeuAlaValSerGlnProLeuLeuSerSerThr
ThrAsnTyrLeuLysAsn...METGlyThrTrpAsnGlySerProThrProTrpSerProCysLysIleAsnLeu
ThrPro...GlnGln...SerPheLysIleGluGluLeu...ThrCys...ProLeuLysGluGlyGluProVal

FIGURE 20.1

48/64

TyrPhe...GlyLysAsnAlaValIleMETLeuIleAsnProGluSerSerLeuArgLysLeuLysLysPheGlu
IleGluTyrAsnValGluGlnArgSerPheGluThrLeuAspProGlyAlaSerSerAlaAsnGlyCysProGly
PheSerProSer...AspLeu...GlnLeu...TyrCysTyrSerSerLeuAspProValSerLeuThrSerLeu
LeuThrLeuSerLeuProGluSerLysLeu...AsnTyrLysTrpSerProArgCysSerProArgLeuArgSer
ThrAlaAspProTrpThrGlyLeuLeuAlaHisAspLeuMETLeuMETThrSerLysAlaProLeuLeuArgLys
SerGlnLeuHisAsnLeuTyrTyrAlaProIleGlnGlnGluAlaValArgAlaValValGlyGlnProProGln
GlnHisLeuGlyPheProValGluMETGly

FIGURE 20.2

49/64

AlaProSerGlyGluGlnArgThrGlyHisTyrProGlyGluAspTrpGlnLeuIleLeuProThrSerProAsn
LeuArgAspPheSerIleTyr...SerGly...IleLeuSerArgValGlyGlnArgProSerProValGlyGln
LysArgProLysArg.....ArgHis...PheMETLys...PheProAspSerAspPheProGluAlaTyrArg
ValThrIleAlaLeuLeuSerArgProGln...ProArgGluTyrProArgArg...ValTyrAspIleThrTyr
ThrAlaProGluGlyHisSerProGlnGlyArgSerArgLys...METLysHisSerLysAspIle...LysSer
LysProArgLysProThrSerHisGlyLeuLeuCysCysLeu...Pro...LysGluSerAlaThrPheProLys
LysGlnAspLeuAlaHisThrLysCysCysMETGluGlyProSer...ProMETThrLeuCysLeuThrGlnAsp
SerGlnLeuSerCysArgHisHisLeuLeuSerGlnIleSerThrSerSer...AsnIleThrArgAsnLeuSer
LeuArgArgGlyLysArgThrIleProProLeu...HisGlyIleSerGlnValProSerLeu...PheProIle
ProArgTyrIleLeuGlyArgThrLeuProSerHisPheIleTyrProAsnCysGly...SerGlyTrpSerGly
ValLeuAspThrSerHisLeuSerGlnIleLeuAspThrAlaLysGlyThr...LysSerArgArgGlnArg...
LeuPheLeu...ThrSerArgGlyPheAlaProAlaLeuGlnThrThrThrArgArgLysValThrLysIleIle
AsnProHisGlyProProLeuSerTyrPheSerLeuTyrCysSerPheThrLeuPheHisSerHisCysThrPro
SerMETProLeuTyrAspGln...LeuProLeuProArgValSerMETGluAsnAlaAlaSerArgLysTyr...
CysProIleVal...GluSerPhe...GlyAsnProHisLeuHisCysProHisProTyrAlaProGlnLeuLeu
SerLeuCysHisSerLeuHisAlaCysLysTyrSerLeuLeuAspArgLysAsnAsp...Ser...LeuSerTrp
ArgThrTrpSerHisCysLeuLeuAspLeuLeuHisProAsnTrpTyrVal...TrpGlyTrpSerSerArgSer
GlyLysArgLysThrCysLysArgSerAsnLeuProThrHisProGlyThrTrpHisLeu...ProLeuGlnArg
ThrArgSerLeuLysThrThr...AsnProProTyrProTyrSerProGlyLysProIle...TyrHisProHis
TrpAlaPro...GlyLeuGlyProLysProTyr...LeuLeuAspMETProProProGluLeuGlnAlaIleCys
PheAsnProCysThr...ThrMETGluGlnLeuGlnHisArgAsnLysHisHisPheArgPheSerArgThrSer
CysPheGlnSerGlyAsnAsnProTyrLeuLysProHisLeuCysLysIle...GlnTyrTyrIleHisAsnGln
LeuProMETHisGlnValGlyAsnSerSerHisThrAsnSerLeuProThrLeuArgAsnIlePheCysLeuTrp
TyrLeuSerLeuSerLeuPheGluTrpLeuPheArgIleTyrValLeuProLeuIleLeuSerAlaProTyrAsp
HisLeuHis...ThrArgPheIleGlnLeuCysHisIle...AlaProGlnGlnLysSerThrHisSerSerPhe
CysTyrArgSerArgSerAlaArgCysThrArgTyrTrpHisTrpArgTyrHisAsnLeuTyrSerValLeuLeu
GlnThrIleSerArgThrLysTrpGlyHisGlyThrGlyArgArgLeuProGlyHisLeuAlaArgSerThr...
LeuProSerSerSerSerProSerLysSerLysSerPheArgLeuAlaAsnArg...LysArgGlyAsnLeuPhe

FIGURE 21.1

50/64

IlePheArgGlyArgMETLeuLeuLeuCys...SerIleArgAsnArgHis...GluSer...ArgAsnSerArg
SerAsnThrThr...SerArgGlyAlaSerLysHisTrpThrLeuGlyProProGlnProMETAspAlaLeuAsp
SerProLeuLeuArgThrSerSerSerTyrAsnIleAlaThrProLeuTrpThrLeuTyrLeu...ProProCys
...LeuCysLeuPheGlnAsnArgSerCysLysThrThrAsnGlyAlaGlnAspAlaValGlnAsp...AspLeu
ProGlnThrProGlyProAlaCys...ProThrIle...Cys.....HisGlnArgHisProSer...GlyAsn
LeuSerCysThrThrSerThrThrProGlnPheSerArgLysGlnLeuGluArgSerSerAlaAsnLeuProAsn
SerThr...ValPheLeuLeuArgTrpGly

FIGURE 21.2

51/64

TTGGTCTTAAGAACACAAATGATATGGCTCCAATGACTGGAGGAACACCAGGGTCCTTGG
TCTCACGCTGATTTAGATAAAACGACTGTCAGGCCTCTGAGCCCAAGCTAAGCCATCCTC
CCCTGTGACCTGCACGTATACATCCAGATGGCCTGAAGTAACCAAAGAATCACAAAAGCA
GTGAAAATGGCCTGTTCTGCTTAACTGATGACATTCCACCATTGTGATTGTTCTCTGC
CCCATCTTAACTGAGCGATTAACCTTGTGAAATTCCTTCTCCTGGCTCAAAACCTCCCC
ACTGAGCACCTTGTGACCCCCGCCCTGCCCCCTAAGAGAAAACCCCTTTGATTATAATT
TTCCACTACCCACCCAAATCCTATAAAATGGCCCCACCCCTATCTCCCTTCGCTGACTCC
TTTTTCGGACTCAGCCCGCTGCACCCAGGTGAAATAAACAGCCTTGTTGCTCACACAAA
GCCTGTTTGGTGGACTCTCTTCACACGGACGCTCATGACATTGGTGCCAAAACCTGGGA
TAGGAGGACTCCTTCAGGAGACCAGTCCCCTGTCTTGCCCTCACTCTGTGAGGACATCC
ACCTACAACCTTGGGTCTCAGACCAACCAGCCCAAGGAACAGCTCACCAATTTCAAATC
AGGTAAGCAGTCTTTTCACTCTCTTCTCCAGCCTCTCTTGCTACCCCTCAAACCTCCCTCT
CTCACTACCCCTCAATCTCCCTGTCTTCCAATTCAGTTCTTTTTCATCTCTAGTAGAG
ACAAAGGAGACACATTTTATCCATGGACCCAAAACCTCCAGCACCAGTCACGGACTTGGA
AGACAGTCTTCCCTTGGTGTTAATCACTGCGGGACGCTGCCTGATTATTCACCCACA
CTCCATTGGTGTCTGATCACGGTGGGGACACCTGCCTTGGTCACTCACCCACATTCCCTT
GGTGGTACGTCAACTGCAAAAGCAGGGGACGCTGCTTTGGCTGCTCACCCACCCCTTC
TCTGTGTCTCTACCTTTCTCTTTAACTTACCTCCTTCACTATGGGCAAACTTCTGCCCT
CCATTCCCCCTTCTTCTCCCTTAGCCTGTGTCTTAAAAACCTAAAACCTCTTCAACTCA
CACCTGACCTAAAACCTAAATGCCTTATTTTCTTCTGCAACACTGCGTGGCTGCAGTACA
AACTTGATAATAGCTTTAAATGGCCAGAATATGGCACTTTCAATTTCTCCATCCTACAAG
ATCTAGATAATTTTTGTGGAAAAATGGAAAAATGGTCTGAGATGCCTGACGTCCAGGCAT
FTTTTTACACATTGGTCCCTCCCTAGTCTCTGCTCCCAATGCGACTCATCCCAAATCTTT
CTTCTTTCTCTCCTGTCTGTTCCCTTCAGTCTCCACCCCAAGCTCTGAGTCCTTTGAATCC
TCCTTTGCTACAGACCCATCTGAACCTCTCCCCTCCTCCCCAGGCTGCTCCTCACCAGGCC
GAGCCAGGTCCCAATTCTTCTCAGCCTCTGCTCCCCCACCCTATAATCCTTTTATCACC
TCCTCTCCTCACACTCAGTCCGGCTTACAGTTTCTGTTCTGTGACTAGCCCTCCCCATCT
GCCCAACAATTTCTTCTTAAAGAGGTGGCTGGAGCTAAAGGCATAGTCAAGGTTAATGCT
CCTTTTTCTTTATCTGACCTCTCCCAAATCAGTTAGCGTTTACGCTCTTTTTTCATCAAAT
ATAAAAACCCAGCCAGTTTCATGGCCCATCTGGCAACAACCCCTTACAGGCTTTACAGCCCT
AGACCCTGAAGGGTCAGAAGGCCGTCTTATTCTCAATATGCATTTTATTACCCAATCCGC
TCCCAACATTAAATAAAGCTCCAAAAATTAAATTCCTGGCCCTCAAACCCACAACAGGAC
TTAATTAACCTCACTTCAAGGTGTACAAGAATAGAGTAGAGGCAGCCAAGTAGCAACGTA
TTTGAGTTGCAATTCCTTGCTCAACTCTGAGAGAAACCCAGCCACATCTCCAGCAAAC
AAGAACTTCAAAACACCTGAACCTGCAGCAGCCAGGCGTTCTCCAGGACCACCTCCCCCA
GGATCTTGCTTCAAGTGCCGGAATCTGACCATTTGGGCAAGGAATGCCTGCAGCCAGG
ATTCCTCCTAAGCCACGTCCCATTTGTGCAGGACCCCACTGGAAATCGGACTGTCCAAT
CACCCGGCAGCCAATCCCAGAGCCCCTGGAACCTCTGGCCCAAGGCTCTCTGACTGACTCC
TTCCCAGATCTTCTCGGCTTAGCAGCTGAAGACTGACACTGCCCAGTCACTTCAGAAGTC
CCCTGGACCATCACGGATACTGAGCTTCAGGTAACCTCTCACAGTGGAGGCTAAGTCCATC
CCCTGTTTAAATCGATACAGGGGCTACCCACTCCACATCACCTTCTTTTCAAGGGCCTGTT
TCCCTTTCCCCCATAACTGTTGTGGGTATTGACGGCCAAGCTTCAAACCCCTTAAACT
CCCCACTCTGGTGCCAACTTGGACAACATTCTTTTATGCACTCTTTTTTCAGTTATCCTC
ACCTGCCCAGTTCCTTATTAGGCCGAGACATTTTAAACCAATTATCTGCTTCCCCGACT
ATTCCTGGGCTACAGCCACATCTCCTTGCCGCCCTTCTTCCAACCCAAAGCCTCCTTCA
TATCTTCTCTCATATCCCCCACCTTAACCCACAAGTATGGGACACCTCTACTCCCTCC
CTGGCAACCGATCACACGCCCATTAATCTCCATTAACCTAATCACCTTACCCTGCT
CAATGCCAGTATCCCATACCACAACAGGCTTTAAAGGGATTGAAGCCTGTTACTTGC

FIGURE 22.1

52/64

CTGCTACAGCACGGGCTTCTAAAACCTATAAACTCTCCATACAATTCCCCCATTTTACCT
 GTCTAAAAACCAGATAAGTCTTACAGGTTAGTTCAGAATCTGCACCTTATCAACCAAATT
 GTTTTGCCTATCCACCCTGTAGCACCCAACTCGTACACTCTTTTGTCTCAATGCCTTCC
 CCCACAACCTCACTATTCGGTTCTTGATCTTAAAGATGCTTTTTTCACTATTCCTTGCAC
 CCTCATCCCAGCCTCTCTTTGCTTTTACCTGGACTGACCCTGACACCCATCAGTCCCAG
 CAGCTTACCTGGGCTGTACTGCCGCAAGGCTTCAGGGACAGCCCTCATTACTTCAGCCAA
 GCTCTTTCTCATGATTTACTTTCTTTCCACCTCTCTGCTTCTCACCTTATTCAATATATT
 GATGACCTTCTACTTTGTAGCCCCCTCTTTAAATCTTCTCAACAAGACACCCTCCTGCTC
 CTTCAACATTTGTTCTCCAAAGGATATCGGGTATCCCCCTCCAAAGCTCAAATTTCTTCT
 CCATCTGTTACATACCTCGGCATAATTCTTCATGAAAACACATGTGCTCTCCCTGCCAAT
 TGCGTCTCCAACCTGATCTCTCAATCCCAACCTCTTCTACAAAACAACAACCTCCTTTCCC
 TCCTAGGCATGGTTGGATACTTTTGCCTTTGGATACCTGGTTTTTGCCATCCTAACAAAAT
 CATTATATAAACTCACAAAAGGAAACCTAGCTGACCCCATAGATTCTAAATCCTTTCCCC
 ACTCCTCTTTCCATTCTTGAAGACAGCTTTAGAGACTGCTCCCACACTAGCTCTCCCTG
 TCTCATCCCCAACCTTTTTATTACACACAGCCGAAGTGCAGGGCTGTGCAGTCCGAATTC
 TTACACAAGGACCGGGACCATGCCCTGTAGCCTTTTGTCCAAACAACCTTGACCTTACTG
 TTTTAGGCTCGCCATCATGTCTCCATGCGGTAGCTCCGCTGCCCTAATACTTTTAGAGG
 CCTCAAAATCACAACTATGCTCAACTCACTCTCTACAGCTCTCACAACTTCCAAAATC
 TATTTTCTTTCTCACACCTGACGCATATACTTTCTGCTCCCCGGCTCCTTCAGCTGTATT
 CACTCTTTGTTGAGTCTCCACAATTACCATTCTTCCTGGCCCAGACTTCAATCTGGCCT
 CCCACATTATTCTGGATACCACACCTGACCCTGATGATTGTATGTCTCTGATCTACCTGA
 CATTACCCCCATTTCCCCCATATTTCTTCTTTTCTGTTCCCTCATGTTGATCACATTTGGT
 TTAGTGACGGCAGTTCCACCAGGCCTGATCGCCACTCACCAGCAAAGGCAGGCTATGCTA
 TAGAATCTTCCACATCCATCATTGAGGCTACTGCTCTGCCCCCTCCACTACCTCTCAGC
 AAGCCGAACCTGATTGCCTTAACTCGGGCCTTCACTCTTGCAAAGGGACTACACGTCAATA
 TTTATACTGACTCTAAATATGCCTTCCATATCTTGCAACCACCATGCTGTTATATGGGCTG
 AAAGAGGTTTCTCTACTACGCAAGGGTCTCCATCATTAAATGCCTCTTTAATAAAAACTC
 TTCTCAAGGCTGCTTTACTTCCAAAGGAAGCTGGAGTCACACACTGCAAGGGCCACCAA
 AGGCGTCAGATCCCATTACTCTAGGAAATGCTTATGCTGATAAGGTAGCTAAAGAAGCAC
 CTAGCGTTCCAACTTCTGTCCCTCATGGCCAGTTTTTCTCCTTCCCATCAGTCATTCCCA
 CCTACTCCCCCATTGAAACTTCCGCCTATCAATCTCTTCTCACACAAGGCAAATGGTTCT
 TAGACCAAGGAAAATATCTCCTTCCAGCCTCACAGGCCCATTTCTATTCTGTCTCATTTTC
 ATAACCTCTTCCATGTAGGTTACAAGCCACTAGTCCACCTCTTAGAACCTCTCATTTCCT
 TCCATCGTGGAACATATCCTCAAGGAAATCACTTCTCAGTGTTCATCTGCTATTCTAC
 TACCCCTCAGGGATTGTTTCAAGCCCCCTCCCTCCCTACACATCAAGCTCGGGGATTGTC
 CCCTGCCCAGGACTGGCAAATTGACTTTACTCACATGCCCTGAGTCAGGAAACTAAAATA
 CCTCTTGGTCTGGGTAGACACTGTCACTGGATGGGTAGAGGCCTTTCCCACAGGGTCTGA
 GAAGGCCACTGCAGTCATTTCTTCCCTTCTGTCTCAGACATAATTCTTGGGTTGGCCTTCC
 CACCTCTATACAGTCCAATAACGGAGCAGCCTTTATTAGTCAAATCACCTGAGCAGTTTTT
 TCAGGCTCTTGGTATTTCAGTGAACCTTCGTACCCCTTACTGTCTCAATCTTCAGGAAA
 GGTAGAATGGACTAATGGTCTTTTAAAAACACACCCCAACCAAACCTCAGCCTCCAACCTAA
 AAAGGAGGATAGAGCCCCAAAACCTCGCAACCAAGCTAGTAATTATGCTGAACCCCTTGG
 GCACTCTCTAATTGGATGTCTTAGGTCTCCCAAATCTTAGTCTTTAATATCTGTTTTT
 CTCCTTCTCTTATTCGGACCTTGTGTCTTCCGTTTAGTTTTTCAATTCATACAAAACCGC
 ATCCAGGCCATACCAATCGTTCTATACAATAAATGCTCCTTCTAACAACCCCAATAT
 CGCCCTTACCACAAAATCTTCTTCACTTAATCTCTCCACTCTAGGTTCCCATGCCG
 CCCATAATCCCTCTCGAAGCAGCCCTGAGAAACATAGCCCATTTATCTCTCCATACCACCC
 CCAAAATTTTTGCTGCCCAACACTTCAACACTATTTTACATTATTTTTCTTATTAATAT

FIGURE 22.2

53/64

AAGAAGACAGCAATGTCAGGCCTCTGAGCCCAAGCCATCATATCCCCTGTGACCTGCACA
TATACATCCAGATGGCCTGAAGTAACTGAAGAATCACAAAAGAAGTGAAAATGGCCTGTT
CCTGCCTTAACCGATGACATTCCACCACTGTGATTGTTCCTGCCCCACCTTAACTGAGC
AATTAACCTTGGGAAATTCTTCTCCTGGCTCAAAACCTCCCCACTGAGCACCTTGTGA
CCCCTGCCCCTCCACTACCCACCCAAATCCTATAAAATGGCCCCACCCCATCTCCCTTAG
CTGACTCCTTTTTTGGACTCAGCCCGCCTGCACCCAGGTGAAATAAACAGCCTTGTGCT
CACACAAAGCCTGTTTGGTGGACTCTCTTCACAGGGACGGGGGTGACAACAACACGGACA
CACATGGAGTGGTTTTAAGGAGCAGAGAGTTTAATACGCAAAAAGAAGGAAGAGGCTCC
CCTGTACAGACACAGAGGGAGGGGGCTCCAAGCCGAGAGAAGGAAACCCCATGTGCAGTG
GAAAAGTGGTTGATTATACTGGGAGGCTGGAGGAGGCGGTGTCTGATTGTCACAGGGCCC
AGGGGATTGGGTTGACCAGGTGTATCATTATGTACCCCGCAAAAACCTGGCCCTCCCA
CCTCAGCCCTTTAATATGCAATGTGGGTTGCCATGATGTTCTGAAAACACATGAATTAT
CTGGAGGGGGCCATGACACTTGGTACATGTGCTGACAAGAAGAGGGTGGGAATCGCCATG
GTGGCCATGTTGGGTGGACCTAGTTTTTTAATAGCCTGCATTTGCATATCAAAGTTTGCTG
GCCTGGCTCTTTAAGCTGTCTTTTCTGTTAGAAAAGGAATGGTTTGGAATGGGTGAGGGT
TGCTTCTTATTACAAGAAATTTCCAAAACCTTTACTCTTTCTAGCTGCCAAAAAACTA
TTTCTTAATAACTTATGTATTACCATAATTAGGCAGCACCAAAGATCCCTGCAGGTCAGA
CCACTGCAATTAACATGCTGGCTTTACTGCTGATTATGGTAGCTGCATCCACCTAGCCTC
TCATATTGCAACTGCCTGACCTCTGCCACCCACGAGCCACTTATCCCCACTTATAATCA
GCCCATTTGATTGTAACATCTGCCACTTATTCCCGACGTTGTGGTATATCCTATAGATG
AATTCATTCAACATCCATTCCAACACCACCTCTCTTGCCTTCTTATACTCTCTGGAGAGT
GAATTACTGAGTCACATGATCTTCACTGCAGTCATTTGTGGCTATGTGACATAGTTCTGG
ACAGTGAACATAGACAGAAGTCCCTGGGGCGGGCTTCCTTTCTGGGATGAGGGCAAAACG

FIGURE 22.3

54/64

GATCTCTTGATCCCAGGAGGTCAAGGCTGCAATGAGCTAAGATCAAGCCACTGCATTCCA
 GCCTGAGTGATAGTGGGAGACCTTGTCTTTAAAACACACACACACACACACACACG
 AGGGCCTTTGACCCTCTTGAGTAGAAGACTCGAGAAGAACAAGTAGAAGGCCAGAGAA
 GAACAAAGTTACTTGAAAGATCTCTTATTAAAGAGAATGTACAAGCTATGAAAAAAAAA
 AACACACACACACACACAAACCTCATCTGGAATGAAAAAACATAATGCATTTGGTTTCT
 GGTTCCTTAGGCTGTTATGGAACAACCAAAGAACATTATTTTGGTTTCTGAGGTCAGAAC
 TATTTTATTTCCCTCAAGCACACTATGCTTATGGTTTGGGGAGAATGAGAAATAGGAAA
 CTAGGAACAGGCTGAAATGGTCTAATCTTGACCATCTAATTCTGCAGTGTCTTATTCTCA
 TTCTAAAAGAGAATGGTTATATTCTGCTGTTCTAGCATAAAAAGTAATGATAAAAAATAAAA
 GATCCCGTATTACCAGACAATAATCCCTTAGACTGTTTTAATGCTTGGTTGAGTATTTGC
 TTATGATCTCAGACTTTAAAAAGATGGTCTCCCCCTATGGTGAAGCTTGTAAATTATGTAG
 GCATCATTAAATGTCTGTTTACTTATCAAAATTTTATCATTGTTAGTTGTATTACTACTTG
 ACAGTCCAATTTATTTAATTGAAAAGATTGGTTAACATTTTATAGTCAAAGTAATTGTTT
 CCTGTGTTTTTCTGTGTTAGGTTATTGGAGTGATGAGTAAAGAATACATACCAAAGGCC
ACACGTTTTGGACCCCTAATAGGTGAAATCTACACCAATGACACAGTTCCTAAGAACGCC
AACAGGAAATATTTTTGGAGGGTAAGTAAGGGAAATTTCTTCAGACCCATTAAATGTTAG
 GAAAAAATGGAGCTAAAAGAGCTGGGTGGCTCACCTTCTCATCCTGTGCTGAGAAATGC
 TGGGGCTCACCCATAAGTATCCAGCATCCCCATGGACACAGGGAATTCTGAACAAATGTG
 ATGAAACCGATGAAATGTCTGGCCTGTAGGTGGTTAGTGATGGAGATACGGGCTATATGT
 GAATCTTGATTTTTGCAATTCATTAGAGCTTTGTAAAGGAAACAGTTTGTGCTTG
 CTTTAAAGGATAGGTTTCAATTTGCATTTCTCCGCAAGGAAGTAGTAATGAGTTACCAAGCCT
 TAGATTTTACCCCTTTTTGATTTCTTGCTGACTTAACTTTAATTGAATGGAAGAGTTATC
 ACAAATGAATTATCTTTTTGGTTTTTTTTTTTTTTTGAGATGGAGTCTCACTCTGTCACCAG
 GCTGGAGTGCAATGGCATGATCTCGGCTCACTGCAACCTCCGCTCCAGGTTCAAGCAA
 TTGTCTGCCTCAGCTCCCGAGTAGCTGGGACTAAGGTGCGCGCCACCATGCCAGTTA
 ATTTTTGTATTTTGTAGTAGAGCGGGTTCCACTATGTTGGCCATGATGGTCTCGATCTC
 TGGACCTCGTGATCCGCCACCTTGGCTCCCAAAGTGCTGGAATTACAGGCAAGAGCCA
 CCGCGCCCAGCCAGGAATGACAAATGAATTACCTTATAAGTAAATGCCATTAAGGAAGGA
 TAGCTGGAAGATGGGTTGAGGGGAATGGAGGACCACAGAACTAGTCCTATTTAAATACAT
 GTGCATGGTAAAAATGATTCCATTTGACAATAGGTTAATTATCTCATAGCATAAGGAAAAAT
 GCTTAACAGTCATATGCAAGATGATAAGCTTTCTATAGCATCCAACCAAAAGATCTAGC
 CAGTACAATTTCTTTGCTATATTAGGGTTAGAAAGGCCCCCAGAGGTGAACCAATTAGA
 TGAATCCTTGAATAAAACACTGGATTAGCAGTGAACAGAAAAAAGTCAGATTGCTTTCC
 TTCTTCCCATAGATGTCTCAGGGATATTTAGTTTCTCAGAAGATAAAGAATTTAGTAAG
 CGTTTTTTTGTGCATACTTACATGAAATGTACATTATTTGAATCTTTAAAAAGAAACAG
 CTGCATGATAACAAAAATGTGTTATGCTTGCTTTAGCTGGTATTTTGCCTAGAACGAT
 TATATCGTTCCGACAAGAAGCTATTCCTAAGAAACAATATTTTAAATCCAGGAAGTTTTT
 CATTTTTAGAAATTTATCTTACTATTTCCCAAGCAAAAGAGGGTAGTTACAGATTCAC
 AGAATCATGTGCTCACAAATTTTATTTAATAATTATTCCTCCTTAAAAATATATTAATCAC
 CTGACTTACAATGGTGGAAACCATGAGTGCAATTTTTCCTTTATTTGTCAATAACGTCTTCT
 CAGAAGTGAGCCACAAAGGTGCATAGTTCTTGGAGTTAAAGGTCTGAATTAAGACAATCC
 AGCATAAGTCTCATTAATGTGTGATTATTTTGGAGAAAAGGCAAGAGTACCTAAGAATCT
 CCCCCCTCACTGTCCAGTTCCCTGTTTCATTTAAAGATTCACTGTAAAGTAACTGAAAGGCT
 TTCTTGGGAGGATTTATTTGAATCAGTCTTTCACATGCAAAGGATATTGTAGAATCT
 CGTTTTTGTGCTGGCAGGAATATGAACATCTGTTGTGAGGAAAGAAAAAGTTTCATGCAAT
 TACACTGCCAAAGAAGGGATGTTCAAGTTGAGAAACCAAGTGACATTTCTGTAACGTGTAC
 TATGAATCAGCGCATTTTAAATCTTCTAGATAATATATGGAAGTGCAAGAGGTGGTAGGA
 AACGGTGTTCATTTTACATATGCGTTATTTTATTTCTGTGTGAGTGACTTCATGGCACC
 CATGCTGTTTTTAAATGAGGATACAGTAAATTCAGTCCGAGGAAGGCTAACTGGAATC
 AACATACCCGTAGCTTTAGAAAAGCAGTTTCCGCACCAGCGAAGAGTACAAGAGCGATGGA
 ACCCATGTTCTTGGAAGTTTGCACATCAGAGTAAACAAACTTGAAAAACCCCTCTTGATA

FIGURE 23.1

55/64

GCAGAATTCACCCAGCCTTGTTCCATTTTCTCTTAACAAAAACACACCGCAAAAGCTCTCA
CAAGCTGCTTTGATGAAGCCACATGTATTTCCCCCTTCACAATTTACAGGAAGTTACTCT
TAAAAGAAAGTGATTCTGGTGTTTACCGCCTGTGTTAAAGGGACAGAGTTCCTTTTATT
TCTGATAACGTTTGAGCGAAATACAGAACTATCTGTAGACTAGCATAGTCGGTACGTGA
GTAAGGAAAAGCAATAACCTGCTGTCCGGTGAGCACAAAATTCCTGCTACGAACAGTGCC
TTACTGCTGCTTGGAGACTGCAAGTCGCAGATCACACTAGGTATTGACTGATTGTATAAG
GAAATTTCTTAAAGTCTAAAGTAAAGGTGGTACCTCCTAAAAAGAGGGGAAGAGAGAAAA
CTTTGTGTGGAAGGATAAGGAGTGTGTTTATAGTTTCAGTAAGAGTGTACGTTTAAATTT
TTCTTCTTCTCTGCCTCTTTGCCAAGTAGCCTGAGTGCATCTGTTATCCAGAAGTAGTA
TTACTCTAGGACAACTTCAAATTCTTCATTCTGCGTTGCCTTTAAGGAACAACATACTT
TCTTCTGTTCTTTTTTCCAAAAACACACGCCTATGGCTCTGTGTGTGGTGTTTTAGCCAG
CCTCCTCCAGATAAGGGGTTCCCTTCCCTCCTTTGCATTGAAAGGAAAGTGCAAGTCTG
GACATGTTTATCAAGAGGAAAAGTGACTTCTCAGTAATAGACTGTCAAATTCGGGCTGCT
GCCCGAGTGTTTCGCTTTGTTATGGCAGGTGAAGTTCACCTTTGCCCCACCCAGTGTTTCC
ACAAAAAGGCAAGGTTCCAAGTATTCATATGAACAAGTGTTACTTTAGGACTTGAGGGT
TGGGGGTGGAGGATGTTTGCATAGTTGAAGCCTTGGGCGGGGGTGTAGGAAACGGCGAGT
ACAGAGGCCATAGAAAAAGCTAAGACTCAGTTTGACGTCGTCAGCCGGCTTGGTCTTCTA
CCCAGTGACTCAAAGCACTAAAAGTCAGCATAATCGGAAGTGAAGTCAGTAGCATCGCCC
ATTTGCCATTCACTGCAGTAGCAAAAGTAGTACTCTGTGGTGGGTAAATCGGTTTGAGGC
AGTCCCTTAAATGAACATTTGTGTTTCATTTTCTGTTATTTTCCCGAACATGAAAAGAC
GATAAAACTGAAATGGAAAAGGTAACGTGACAAAAGTGTGCCTTACCTGTTTTCCGCCCTGA
TTTCTGCTGATTCAAGACTATTCTGGCTAAACTGATTGGATTCTTTTTCTAACTAGGCAG
TAGGGGATCAGAAATCACACACGGTACCGGCTGTGTTTATTCTGAGAGGTGCTGGGGAGC
TTTGGGTCTGACTTCCTTTTACATGCCTGTCTTCTTTTTGGACAGATCTATTCCAGAGG
GGAGCTTCACCACTTCATTGACGGCTTTAATGAAGAGAAAAGCAACTGGATGCGCTATGT
GAATCCAGCACACTCTCCCCGGGAGCAAAACCTGGCTGCGTGTGAGAACGGGATGAACAT
CTACTTCTACACCATTAAAGCCCATCCCTGCCAACCAGGAACCTCTTGTGTGGTATTGTGCG
GGACTTTGCAGAAAGGCTTCACTACCCTTATCCCGGAGAGCTGACAATGATGAATCTCAG
TAAGTGGATTACAGAACAAAAAATAAAAAATGCCAGTAATGTCGGTCTGCCCCCTTGA
ACTAATAACATGTTGTTTAAATTATACGGCTTTGTGTCATGTGTTGGATGAAGTAGGTGGCTT
AAGCTAGGGACTAGGAAGAGGAAAAACATTTTTTTGAGTCCCTATTAACATTAGGAAACT
TGATCATTTAAAAGTATATATATATATATAGAGGAGCTACCTTGAGTTTTGAATTCAGGATGT
TACAGGAAGAAATATATGTCCAATTCTAATTTATCCAAAAGCAGTTGGGAGAATTACAGG
GATTGGTCCAGACATGCTGCGTATGCAAGGTATAGCCCTCATCTGTGGTACTTTGGCAGG
GCTTAGACTGCATCAAAATATTTATAGATGTACATTTGAGTGTACAGTTAGGATCTGATG
TGGAACATTGTAAGATCATTGCTAGAAAACTTTGTCATAATTTTTCAATATTATTCTAA
GTGAATAACCGTAAAGATTTTACATCTTAGCTTCCCTTACAGTAAAAAACTATCTG
ATCTCTTGATCAGTATTATAGTAGCCACCTATCACTTTATCTTAACAAATTCTCAATTCC
TTAGGTTTTATGTGCTTTTACTTCTTTTATTGATTAAATTTGCTGTGATGACCTCTCTCT
GCAGAGGGCTGCATCATTTTGGTCATTCTCAAGTGATCTCTTTGAGCAATTTAAGAATTG
CCATAAGATTCTAACCTCTGCTGTAACCTATGGTTGTGTGTTCTTGGTTAGACCACTAAAT
CTTATTAGCAGTTTTTAAAAATTATTCCTTTTGGTTTAGAAGTTAAGACTAAATGCTGAAG
TTTTTGTAACTTTTGGTTTTGATATCATTTCAAACCTTAAGAAAACATTTGAAGAAAAGGA
CAAAGAATTTCCACTTACCCTTTACCCAGGTTTACCAGTTATTGATAAGTATATCCATTT
GCTTTACCAGAAGGCTAACTTGTTTTAGTTCTCATTTTACCTTTGAGACATTTGGAATA
AATATCAATGTTAACATAAATTGGAATTTTGACTTTGATTTTAGGACCAATGAACAAGCC
AAGTACTTACCCTAGTCATATATAATCCAAGTGTATGGTTATTTGGTATTCAATCCACAC
TTCATTTTACTTGATCTCCCTTAAGATTGCAAGATTGTGTTTGCAGTTTTTCTGAAAATC
TGGGGCTATAAAAGCATCAGGACCTCCCCCGTAGGGGAGGTGCTGTGTTTGGGGTCTTCA
CACAACAGGTTACCCTTGAGCTTCAGGAAAAGAACTGGCTCTCAGTTCCCGAGTTCCAGC
TTAATGGGTCTAATTAGGTCTGACCAAAAAGGTGGCAGTTCTTTTCCCTCATGTCTCTT
CAGCGCTCCCCGAGACTCTGGAGACTCTGTGATATCCCTAGGGCTGAGCCTCCCAGGAAC
CATTCGGCTGTTGTGGCATCTGTGTATGCCATGCCAGTGCTGAGGACCTAGTAACAAAC

FIGURE 23.2

56/64

GACAAATGCACAGGCACAGTGGCATT TTTTGTGGAAC TCGTATTCCAGCTGTGCGTCTCAG
AAGAAGCGCACAGCTCCCTCCTGGCTTTCTTAACATAGTGAGCCACTTCCACTTAAGGGT
CTCCTTACATTCCCTTGAGTTTAATCATT CATGGATT CAGAGGAAAGTCTTTTGATTTTG
CTTTTCTTTAAACAGTTCATTTGAGGTGACCTACCCAGTGA CTTTGCACCAACCACAA
GAAACTTTTTTGCATGCTTCCCGCACCCCTGTGCCAATCAAGGGAAGGGTTTAAAGGCCTG
GCGTTTTTATTCCCTCAAAGAAAGTTTTTGCACAGTATTTTAAGGTTCAAGTGCTTCTACT
TTGTGTTT CAGAAGCAACTGT CATATATACTGTGAAATGACACCTTTTATTTATCCCTTTT
TATTTATGCAGTATGTCCCCTTTTATTTTGGCAGAATTTTTTCTAAATGGTGGTTTAAACA
TTTTCAAGCACATTTTCATTGTCCAATATT CATAGTAAAGAATGAGAGTTAACAATAACCA
GTCACATTAAACAAGATTCCCTGCTGCCAGTTGTGAAACCGGTTGTCTTAGGCGTGGCAG
CTGATGATTGAGACTGTGATCAGGAAAATTTCCACTATTTTCATCAGGCCTAATAGGTAGA
TTGTGTCTCCAAATGAACTGTGTTGGGTTTCCATGCTTAAAGCACAAATAGAGGTGGTGCA
AGAATCTCCATGAGGGCTTAAATGGCAGTGATGGTT CAGGCGGTAGAGTTTGGAGAAGAA
GGGATTTGAAAACAAACCAAGGAAAAGTAAGTAGCCAGAAATCACAAAATGGCATT
TTTCTAAAAACAAAGGAAAAAGGAATAAAAGAACTAATAAGTTTGAAACCCCTACCCCTCC
CAAATTTGGCAGGGGGGAGGTATTTTTTTTCTATCTATCTAACTAACCCATCTAGAAAA
CAGTTGACCAAAATTATAGACTTCTAAATGTAACTCTGCTTTCTCAGTTTCAGTTGAAAAAG
AGACTTTGTTTTGCCTACTGCAGAACTTCTAGGTTCTTTCTTATAGTCTTGGGGTTCTTA
TTATAGATCGAAAATGTGAGTCGGCATAATTAAGCCATTTCGGAGTCTTCAGAAGCAGTTC
ACTCTTGAAATGACTCCGTCGCCCTACAGCCATTTAAGATTT CAGAACAAAAACAGATCT
TGATTTTCTTTTTCATGTAACTCAAGCTGTTGCTGAGTGGGAGAGTCAGAAATGACACC
AGCTCCACTGATTACTCAGCTGCTGAAGGATGATTTTTTAAATGCACCTTTACTGTATA
TGGACTTCCTAATTTCCACCTGTAGAGCATCTTAGGGAGGCTAACATGTCACTCTGGATG
TTCTTTTAGAATAAGATGCAAATCTATTTTTCTGAAGGCATTAGAGATAGCAAACATTTA
TTGTGAGTTTACTATATACTAGGCACTGTGCTAAGTGTTTTGCATAGAAAGTTTAAATTT
CTGGCTTTTTTGTGGCCCAATCATAAGTTTCATATCAGTTCAACATTCAAATTATATTA
AGGTACTTAAGAAGAATCCCTGGCTAAATGTGAGGGGCAGTGCCACAGATGGACTGAAAC
TTTATGCTTATTGCACATTTATGCTATTATTATTTGTTGAATTATAGAACCAAGGGAGTG
TGGAAGCCACTGGAAAAAATATGAGACTTAGATACATAATTTGAGTAAAAATGGCTCAAA
GTCATGAGGGTAAAGTTTTTTGTATTTCCATTTTATTTCGAGCGGCATCGTTTTTAAAT
CATTATGAATTTGACCTATATAGATGTTTCCAAATAATTCTTTTTCACCTTCATAAAAT
TCCTTCTGTGGCTGTGAGATGCCTTGCTATCAGTTTTCAAGCTTAGTTGTCTTTCTCA
TCCTTTACCATTTTAGCTTTTAAAAAACAAAGTGACAATTAGAACTTCCTGCCTGCTGGG
CCTCACTGAAAGACCGATATTGGCCTGATAAGGAGATATTTATTTTGTTTTAGTGGCTTC
AGAAATCCCTCTCCCTCAGCAAGCTTTCATCACGGCCCCCCCCGTCAGCATCTTCCCTGA
TAGCGTTCTTCTCTGTGTTTATTCTGGGGCTTCAGGCTCGCCAGGAGGAACTGATAACC
GCTGGCAGGAGATAACATTCTCTAAGGGGCTCTCAAATTGGAATCGAATCCCTCAAGCCA
GTCAGCCTAGAGAATAACATTTAAAGGGTTCAGTTCTGGAGTTTCACAGAGTTCATTCTA
GACCTATCAGATAGCAAGTGTGGAGTTCTTTCTCAACTAAATTCAGCAGAGACATTTT
TAGACGATGAAGGATATTTGCACAAAGGCTTCAGCATGATCCCCCAAACCTGCTGCCTCT
GAAGGCATCTCCACACATTGACAGCCAATGCCTTCAGTGCGTTCTTAGGGCAGGTGTCCT
GGCTTGAGTGACTGTCTCCAATAATCAGAGCTCAAACATAACATCGTATGTTTTACTTT
TGGTTTCCAGGCAAGGCTGAGCAGGGAATTTTCAGTTTTCCCTGCCAGATGGGTGTTTT
TTCTGTAAGGCATCATTTATTGTGTAGCGAGGAGACAGGGCTGGCTGTGGCAGGGATAGT
CTAGAAGTGTCTCATTTGCTGTGTTCTTAAATAGTATCTTTACCAAGTAATAACGTGCC
GTCTTTGGGAATAAGTGCTTTCTCTTAGCCTGTCTGTTTTCTTGGGTGCGCTAAGTAA
TTGAACTGGCTCAGGAAGTACCTATTGTGGTTTGGCAGAGGTGACTGTCACGCCTTGTA
CTCCAGGGGCCAGCACTGCTGGGATCCTGGCTAGACCAGACAGAGCCTTGGTGAAGTGCT
TAGGCTGTCTGCACATCGCGAGGAAGGTGATTCACTTCGCTAAGCTCCTTGGCATAGG
CAGTTTGAACAGGGCTTTATCAAATTCGTATTCAACAAGAGTAGAAGCGAAAATTGATGA
CTGTGTATTACTTGAAATGAGTCTTAATCTTTTACATTTAGTTCTCAGGGTATGCTGATT
TCCTTTAGGTAAACCATGAACATCAGAAAGACTTTTATTAACTATGACAGGGTCCCCAC

FIGURE 23.3

57/64

CCCAGTATTTTTCCACTCCATTAAAAATGGAAGTTTTTTTTTTTTTTTTCTTTTTTGAGAC
 AGAGTTTTGCTCTTGTGTTGCCCAGTCTGGAGTGCAATGGCACAATCTCGGCTCACCACAAC
 CTCCACCTCCCAGATTCAAGCGATTCTTCTGCCTCAGCCTCCCAAGTAGCTGGGATTACA
 GGTGTGCGCCACCACGCCCAGCTAATTTTGTATTTTAGTAGAGATGGGGTTTCTCCATG
 TTGGTCAGGCTGGTCTCGAACTTCCGACCTCAGGTGATCCGCCCACCTCGGCCTCCCAA
 GTGCTGGGATTACAGGCAAGAGCCACTGCATCCAGCTTAGGCTATCTTACTCCAGCCTAA
 ACAGCAATTTTCTATCATAAGGTCTGTACTAATGAAAACAGAATCACCCAAGGCTGCTGT
 TTGTTCTGTCTGTGCTGCCATTGTCCGCATTTTGTGCTGAGGAGGAAACGGAACCTGCACTTT
 TGAGTGAGTGGCCAGAGCCTTCTAGAATGAGAGTGCGTTTGAAGCCAGATATGTGGCGA
 TTGTGTCGCCAGCTGTTACTCAGGTTTTCTCAAGAAGGAGGAGCAACTTTGGCAGTTTTG
 CTTCACTTCTCTCTAGCCCTCTGTGTAATCGCCCCCTTTTTCTTTATTTTCAGCACAAACAC
AGAGCAGTCTAAAGCAACCGAGCACTGAGAAAAATGAACTCTGCCCCAAAGAATGTCCCAA
AGAGAGAGTACAGCGTGAAAGAAATCCTAAAAATTGGACTCCAACCCCTCCAAAGGAAAGG
ACCTCTACCGTTCTAACATTTACCCCTCACATCAGAAAAGGACCTCGATGACTTTAGAA
GACGTGGGAGCCCCGAAATGCCCTTCTACCCCTCGGGTCGTTTACCCCATCCGGGCCCTC
TGCCAGAAGACTTTTGAAGCTTCCCTGGCCTACGGGATCGAGAGACCCACGTACATCA
CTCGCTCCCCCATTTTCAATCCTCCACCCTCCAAGCCCCCTCTGCAAGAAGCAGCCCCGACC
AAAGCCTCAAGAGCTCCAGCCCTCACAGCAGCCCTGGGAATACGGTGTCCCCCTGTGGGCC
CCGGCTCTCAAGAGCACCGGGACTCCTACGCTTACTTGAACGCGTCTTACGGCACGGAAG
GTTTGGGCTCCTACCCCTGGCTACGCACCCCTGCCCCACCTCCCGCCAGCTTTTCATCCCCCT
CGTACAACGCTCACTACCCCAAGTTCCCTCTTGCCCCCTACGGCATGAATTGTAATGGCC
TGAGCGCTGTGAGCAGCATGAATGGCATCAACAACCTTTGGCCTCTTCCCGAGGCTGTGCC
CTGTCTACAGCAATCTCCTCGGTGGGGGAGCCTGCCCCACCCATGCTCAACCCCACTT
CTCTCCCGAGCTCGCTGCCCTCAGATGGAGCCCGAGGTTGCTCCAGCCGGAGCATCCCA
GGGAGGTGCTTGTCCCGGCGCCCCACAGTGCCCTTCTCCTTTACCGGGGCCCGCCGACGA
TGAAGGACAAGGCCCTGTAGCCCCACAAGCGGGTCTCCACGCGCGGAACAGCCGCCACGG
CAGAACATGTGGTGACGCCCCAAGCTACCTCAGCAGCGATGGCAGCCCCCAGCAGCGACG
AAGCCATGAATCTCATTTAAAAACAAAAGAAACATGACCGGCTACAAGACCCCTTCCCTACC
CGCTGAAGAAAGCAGAACGGCAAGATCAAGTACGAATGCAACGTTTTCGCGCAAGACTTTTCG
GCCAGCTCTCCAATCTGAAGGTAGGCCCTTGAGAGAGAGCAGTCCAAGGGGCTGTGAGTGC
 ATGCTTGTGTTTGTATTTAGCTTGCTTTCCATGGGGTATCGATTGCATTTGCAGTAGTAT
 GAGCCCCCGGTTGGGGATAGTGGGTATGGATTCGCCTGGCTTTTGCCACTTCTAGCTCT
 TTGACTTTGGACAAGTGAATTAGGTGAAACTCCCTGCTAGCCTGTGATTTTGTGCTTTTAA
 GAAAAACACCATTTCTGAAAACATGAAGATTTCTTCTTTTTTAAGACTGTCTTGATGCTTTT
 CTTAAGATAATTTGTCATCAACACTTGAGTCTTGGAGCAGAAATGTTAGGTCTCAGAGCCAG
 CTTGAGAGCAGAGCTAACACATGTGGCTTCTTCCCAGGTCCACCTGAGAGTGCACAGTGG
AGAACGGCCTTTCAAATGTGAGACTTGCAACAAGGGCTTTACTCAGCTCGCCCCACCTGCA
GAAACACTACCTGGTACACACGGGAGAAAAGCCACATGAATGCCAGGTGCGCAGTATTTT
 CTGGGTAGACCTTCTGACCTTTGTAGAAAATGTCTGTGAGTCAACCTCCCATGTCTTATA
 TAGCCCGTAGTTAAAGCCAACACCAGATTCTGCGTTGTCCCATCCTGGACTGATGGCACT
 ATGGTCCCTCCCAGTACTTTGTATCTGCTGATGACTTGAGATGGCACAGCCAGCTTCCAG
 TGGGTGGGAAAATGGTAGGGGAAATAAACAGCCCCCTCGTGTGCTGTGTGCCACATCCCC
 CCGTTTGTCTAATACCACACTGGAGGTGCCACAAGGAGGCTTCTCACCTCCTAGGTTGCT
 GGGCGTTGGCCGGTAAGCCTGCCCCCTCCCGTTGGCAACTCTTAATCTTCTGGCCTTCTGT
 TCTCCCTTCCCTGCTGTCTCTCTCCCCCTACACTGTAGGCTGCCCACAAGAGATTTAGCAG
CACCAGCAATCTCAAGACCCACCTGCGACTCCATTCTGGAGAGAAACCATACCAATGCAA
GGTGTGCCCTGCCAAGTTTACCCAGTTTGTGCACCTGAACTGCACAAGCGTCTGCACAC
CCGGGAGCGGCCCCACAAGTGTCTCCAGTGCCACAAGAATACATCCATCTCTGTAGCCT
CAAGGTTTACCTGAAAGGGAAGTGCCTGCGGCCCGCGCCTGGGCTGCCCTTGGGAAGA
TCTGACCCGAATCAATGAAGAAATCGAGAAGTTTGACATCAGTGACAATGCTGACCGGCT
CGAGGACGTGGAGGATGACATCAGTGTGATCTCTGTAGTGGAGAAGGAAATCTGGCCGT

FIGURE 23.4

58/64

GGTCAGAAAAGAGAAAAGAAGAACTGGCCTGAAAAGTGTCTTTGCAAAGAAACATGGGGAA
TGGACTCCTCTCCTCAGGGTGCAGCCTTTATGAGTCATCAGATCTACCCCTCATGAAGTT
GCCTCCCAGCAACCCACTACCTCTGGTACCTGTAAAGGTCAAACAAGAAACAGTTGAACC
AATGGATCCTTAAGATTTTCAGAAAACACTTATTTTGTTCCTTAAGTTATGACTGGTGA
GTCAGGGTGCCTGTAGGAAGTGGCTTGTACATAATCCCAGCTCTGCAAAGCTCTCTCGAC
AGCAAATGGTTTTCCCTCACCTCTGGAATTAAGAAGGAACTCCAAAGTTACTGAAATCT
CAGGGCATGAACAAGGCAAAGGCCATATATATATATATATATATCTGTATACATATTA
TATATACTTATTTACACCTGTGTCTATATATTTGCCCTGTGTATTTTGAATATTTGTGT
GGACATGTTTGCATAGCCTTCCCATTAAGACTATTACCTAGTCATAATTATTTTTTC
AATGATAATCCTTCATAATTTATTATACAATTTATCATTAGAAAAGCAATAATAAAAAA
GTTTACAATGACTGGAAAGATTCTTGTAAATTTGAGTATAAATGTATTTTGTCTTGTGG
CCATTCTTTGTAGATAATTTCTGCACATCTGTATAAGTACCTAAGATTTAGTTAAACAAA
TATATGACTTCAGTAACTCTCTCTCTAATAATGGTTTGAAAATGAGGTTTGGGTAAAT
GCCAATGTTGGACAGTTGATGTGTTTCATTCTGGGATCCTATCATTTGAACAGCATTGTGA
CATAACTTGGGGGTATGTGTGCAGGATTACCCAAGAATAACTTAAGTAGAAGAAACAAGA
AAGGGAATCTTGTATATTTTTGTGTAGATTTCATGTTTTTCCCCCAGCCACAATTTTACC
GGAAGGGTGACAGGAAGGCTTTACCAACCTGTCTCTCCCTCCAAAGAGCAGAATCCTCC
CACCGCCCTGCCCTCCCCACCGAGTCCTGTGGCCATTAGAGCGGCCACATGACTTTTGC
ATCCATTGTATTATCAGAAAATGTGAAGAAGAAAAAATGCCATGTTTTAAAACCACTGC
GAAAATTTCCCCAAAGCATAGGTGGCTTTGTGTGTGTGCGATTTGGGGGCTTGAGTCTGG
GTGGTGTTTTGTGTGTGGTTTTTGTGTCTTTTTTTTTTTTTTTTTTAAATGTCAAAT
TGCACAAACATGGTGCTCTACCAGGAAGGATTCGAGGTAGATAGGCTCAGGCCACACTTT
AAAAACAAACACACAAACAACAAAAAACGGGTATTCTAGTCATCTTGGGGTAAAAGCGGG
TAATGAACATTCTATCCCCAACACATCAATTGTATTTTTTCTGTAAAACCTCAGATTTTC
CTCAGTATTTGTGTTTTTACATTTTATGGTTAATTTAATGGAAGATGAAAGGGCATTGCA
AAGTTGTTCAACAACAGTTACCTCATTGAGTGTGTCCAGTAGTGCAGGAAATGATGTCTT
ATCTAATGATTTGCTTCTCTAGAGGAGAAACCGAGTAAATGTGCTCCAGCAAGATAGACT
TTGTGTTATTCTATCTTTTATTCTGCTAAGCCCAAAGATTACATGTTGGTGTCAAAGTG
TAGCAAAAAATGATGTATATTTATAAATCTATTTATACCACTATATCATATGTATATATA
TTTATAACCACTTAAATTGTGAGCCAAGCCATGTAAAAGATCTACTTTTTCTAAGGGCAA
AAAAAAAAAAAAAAAAAAGAACACTCCTTTCTGAGACTTTGCTTAATACTTGGTGACC
TCACAATCACGTCGGTATGATTGGGCACCCTTGCTACTGTAAGAGACCCTAAAACCTTG
GTGCAGTGGTGGGGACCACAAAACAACAGGGAGGAAGAGATACATCATTTTTTTAGTATT
AAGGACCATCTAAGACAGCTCTATTTTTTTTTTGGCACTTTATGATTATGTGGTCACACC
CAAGTCACAGAAATAAAAACTGACTTTACCGCTGCAATTTTTCTGTTTTCTCCTTACT
AAATACTGATACATTACTCCAATCTATTTTATAATTATATTTGACATTTTGTTCACATCA
ACTAATGTTTCACCTGTAGAAGAGAACAATTTTGAATAATCCAGGGAAACCAAGAGCCT
TACTGGTCTTCTGTAACTTCCAAGACTGACAGCTTTTTATGTATCAGTGTGTGATAACA
CAGTCCTTAACTGAAGGTAAACCAAGCATCACGTTGACATTAGACCAATACTTTTGAT
TCCCACTACTCGTTTGTCTTTTCTCCTTTTGTGCTTTCCCATAGTGAGAATTTTTAT
AAAGACTTCTTGCTTCTCTACCATCCATCCTTCTCTTTTCTGCCTCTTACATGTGAATG
TTGAGCCCAATCAACAGTGGTTTTATTTTTTCTCTACTCAAAGTTAAAACCTGACCAA

FIGURE 23.5

FIGURE 24

60/64

GAATTCCGGGAAGCCAGACGGTTAACACAGACAAAGTGCTGCCGTGACACTCGGCCCTCCAGTGTTGCGG
 AGAGGCAAGAGCAGCGACCGCGCACCTGTCCGCCCGGAGCTGGGACGCGCGCCCGGGCGGCCGGACGAAG
 CGAGGAGGGACCGCCGAGGCTGCCCCAAGTGTAACCTCCAGCACTGTGAGGTTTCAGGGATTGGCAGAGG
 GGACCAAGGGGACATGAAAATGGACATGGAGGATGCGGATATGACTCTGTGGACAGAGGCTGAGTTTGAA
 GAGAAGTGTAACATATTGTGAACGACCACCCCTGGGATTCTGGTGCTGATGGCGGTACTTCGGTTTCAGG
 CGGAGGCATCCTTACCAAGGAATCTGCTTTTCAAGTATGCCACCAACAGTGAAAGAGGTTATTGGAGTGAT
 GAGTAAAGAATACATACCAAGGGCACACGTTTTTGGACCCCTAATAGGTGAAATCTACACCAATGACACA
 GTTCCTAAGAACGCCAACAGGAAATATTTTTGGAGGATCTATTCCAGAGGGGAGCTTCACCACTTCATTG
 ACGGCTTTAATGAAGAGAAAAAGCAACTGGATGCGCTATGTGAATCCAGCACACTCTCCCCGGGAGCAAAA
 CCTGGCTGCGTGTGAGAACGGGATGAACATCTACTTCTACACCATTAGGCCATCCCTGCCAACCAAGAA
 CTCTCTGTGTGGTATTGTGCGGACTTTGCAGAAAGGCTTCACTACCCTTATCCCGGAGAGCTGACAATGA
 TGAATCTCACACAAACACAGAGCAGTCTAAAGCAACCGAGCACTGAGAAAAATGAACTCTGCCCAAAGAA
 TGTCCCAAAGAGAGAGATACAGCGTGAAAGAAATCCTAAAATTGGACTCCAACCCCTCCAAAGGAAAGGAC
 CTCTACCGTTCTAACATTTCAACCCCTCAGATCAGAAAAGGACCTCGATGACTTTAGAAGACGTTGGGAGCC
 CCGAAATGCCCTTCTACCCCTCGGGTCTTTACCCCATCCGGGGCCCTCTGCCAGAAGACTTTTTGAAAGC
 TTCCCTGGCCTACGGGATCGAGAGACCCACGTACATCACTCGCTCCCCCATTTCCATCCTCCACCACTCCA
 AGCCCTCTGCAAGAAGCAGCCCCGACCAAGCCTCAAGAGCTCCAGCCCTCACAGCAGCCCTGGGAATA
 CGGTGTCCCTGTGGGCCCCGGCTCTCAAGAGCACCGGGACTCTACGCTTACTTGAACGCGTCTACGG
 CACGGAAGGTTTGGGCTCTACCTGGCTACGCACCCCTGCCACCTCCCGCCAGCTTTCATCCCTCG
 TACAACGCTCACTACCCCAAGTTCCTCTTGCCCCCTACGGCATGAATTGTAATGGCTGAGCGCTGTGA
 GCAGCATGAATGGCATCAACAACCTTTGGCTCTTCCCGAGGCTGTGCCCTGTCTACAGCAATCTCCTCGG
 TGGGGGCGAGCTGCCCAACCCATGCTCAACCCCACTTCTCTCCCGAGCTCGCTGCCCTCAGATGGAGCC
 CGGAGGTTGCTCCAGCCGGAGCATCCCAGGAGGTGCTTGTCCCGGCGCCCCACAGTGCCCTTCTCCTTTA
 CCGGGGCGCGCCAGCATGAAGGACAAGGCCGTGAGCCCCACAAGCGGGTCTCCCACGGCGGGAAACAGC
 CGCCACGGCAGAACATGTGGTGACGCCCCAAGCTACCTCAGCAGCGATGGCAGCCCCCAGCAGCGACGAA
 GCCATGAATCTCATTAACCAAGAAACATGACCGGCTACAAGACCCCTTCCCTACCCGCTGAAGAAGC
 AGAACGGCAAGATCAAGTACGAATGCAACGTTTGCGCCAAGACTTTCGGCCAGCTCTCCAATCTGAAGGT
 CCACCTGAGAGTGCACAGTGGAGAACGGCCTTTCAAATGTCAGACTTGCAACAAGGGCTTTACTCAGCTC
 GCCACCTGCAGAAACACTACCTGGTACACACGGGAGAAAAGCCACATGAATGCCAGGTCTGCCACAAGA
 GATTTAGCAGCACCAGCAATCTCAAGACCCACCTGCGACTCCATTCTGGAGAGAAACCATACCAATGCAA
 GGTGTGCCCTGCCAAGTTCACCCAGTTTGTGCACCTGAAACTGCACAAGCGTCTGCACACCCGGGAGCGG
 CCCCACAAGTGCTCCCAAGTGCCACAAGAACTACATCCATCTCTGTAGCCTCAAGGTTACCTGAAAGGGA
 ACTGCGCTGCGGCCCCGGCGCCTGGGCTGCCCTTGGAAGATCTGACCCGAATCAATGAAGAAATCGAGAA
 GTTTGACATCAGTGACAAATGCTGACCGGCTCGAGGACGTGGAGGATGACATCAGTGTGATCTCTGTAGTG
 GAGAAGGAAATTTCTGGCCGTGGTTCAGAAAAGAGAAAAGAACTGGCCTGAAAGTGTCTTTGCAAGAA
 ACATGGGGAAATGGACTCCTCTCCTCAGGGTGCAGCCTTTATGAGTCATCAGATCTACCCCTCATGAAGTT
 GCCTCCAGCAACCCACTACCTCTGGTACCTGTAAAGGTCAAACAAGAAACAGTTGAACCAATGGATCCT
 TAAGATTTTCAGAAAACACTTATTT

FIGURE 25

61/64

[illegible]

FIGURE 26.1

62/64

GCTACTGCCACCGCCACGGCCACCACCACAACCTACTACCCTACCATTTCACCACATCACCTCTACCATCA
 CTACTGGCCTCATGGATAGCAGTCACCTGGAGATGACGTCCTGGGCGGCTCTGCCCCCTTCTATCCAGCAG
 CAGCACTAATGTCCGGAGACCCCAAGCTCACTTTTGTAGTACTCGGTTTACAATGCTGATTATTACATGCAA
 GAAGCTAAGAAGCTGAAGCACAAGCTGATGCACTGTTTCGAGAAATTTGGCAAAGCTGTGAATTATGCTG
 ATGCCGCCCTCTCCTTCACTGAATGTGGCAATGCCATGGAACGCGACCCTCTGGAAGCAAAGTCCCCATA
 CACCATGTACTCTGAGACTGTGGAGCTCCTCAGGTATGCAATGAGGCTGAAGAAGCTTTGCAAGTCCCTTG
 GCTTCGGATGGGGACAAAAAGCTAGCAGTACTATGCTACCGATGTTTATCACTCCTCTATTTGAGAATGT
 TTAAGCTGAAGAAGGACCATGCTATGAAGTACTCCAGATCACTGATGGAATATTTTAAGCAAAATGCTTC
 AAAAGTCGCACAGATACCCTCTCCATGGGTAAGCAATGGAAGAAGCACTCCATCCCCAGTGTCTCTCAAC
 AACGTCCTCCCCCATCAACGCAATGGGGAAGTGAACAATGGCCAGTCACCATTCCCCAGCGCATTCAAC
 ACATGGCTGCCAGCCACGTCAACATCACTAGCAATGTGTTACGGGGCTATGAACACTGGGATATGGCCGA
 CAAACTGACAAGAGAAAAACAAGAATTCTTTGGTGATCTGGACACGCTGATGGGGCTCTGACCCAGCAC
 AGCAGCATGACCAATCTTGTCCGCTACGTTTCGCAAGGACTGTGTTGGCTGCGCATCGATGCCCATCTTGT
 TGTAGTGGGTGTTCTCAGATCTCTAGCATCACGACCCATCACTCTACCTCTACCAGCGCACTGATGGTCA
 CTGGTGGAAGTCCACTCACTGGGGAACGTTCTCTTTGGTTATGTTTGTGTTTTATGCTTCTTTTGTATCT
 GTAAAAACAGAAGTCATTGTAAGTTGACACTACAACCTTAAGGGCAGTGACGTTTTATTACTTAGTCAT
 TTTTTTCTTTTAGCATTTGATATGCATTTCTCAGATTCACCATCTTTTTGTGCTTTATGGAATGACAG
 TCCCTACAATATTGTTTTAAGCCACACTACCCAAAACAAGAATGGGAAGCACTTGTGATAAAGACAGG
 CTCCTGAGAAATGCAACAAGTGGTCTTACATATACATGAGAACTTAGACACAAGGGACCATCCCCAAAC
 TCTACTCTTATACCCAGAAAAGACATATTTTCAAGATCTGTCAAACCTTTTGTGTATCCACAGATTCAAT
 CTTCAAGGTGAGAATTTTCTTGTCAAAACCCACTGGTTAGATGTTGTAGCAACATCATAAAATCAAGAGT
 ATCAAGAAAATAAATGAGCATAGCAATGCTACTCTTAAAAAGATGCTATGCCACACAACCAGAGGACTTT
 CTTGTTAGCATCCCTTTCTGATTCCCTATTTTGTAAATTTTAAATGATAAGAAGAAAGGGTGACATTTAT
 TTTGACAAGTTTTAGGCATCAGCTGGCATCAGTGTTTTTCAACTCCATTATTTGAAGTGTAATCCTCAC
 CTGGGGTCTCTGTGTGCAAAGCTGTCTTTTTGAAGAAGAGTTTGGTTGATGCATGCCTTAGTAGCCAAA
 ATGCTACACTCTAGACTTACAAGTGGGAGTTAAGAGAGGTCCTGGAAGTGTCCAACAAGGAATTCACACC
 TCTGCCTCCTTTGCAACAACAACATTTACACAGTTGGTAAGTGGGTCCATAACTGGCAGGATTTTTAAAT
 TGTATTTTGTCTCAAATCTATGGGAACAAAAGTCAAGGTATCACTACCTAGAAGTAATGATATACAGTTTT
 CTTCTAGTGGCTTGAAAATCTGGACTTCTCAATTAATTATTCACATTTTCTCTCTTATAGGTTTTCTGT
 TTTCTACTTTCTTTTTTCTCTTATCTGTGTTTTCCCTTTCTCTTTGTTGGCTCATTAACTTTTGACTGAAT
 TACAATTACTCCTTTTATTAAAGTCCATATTTTGTGAATCATTTCCATGAAAATTTCTAAGAAAACTCC
 AAACCTCTAAATAGTACTAACTTTTATTTTAAAAAGATGAGTCGTGGGGTAGTGCTTCACTTGAGAT
 GCTTTGAAAGAGCCCTAAACATTGGGAACCATTCACCTAATTTGGAGACATTTCTCACTGGTTGTGACTA
 CCCCCCTTATGATCCTTCACATTCATTTTATGTCCCTAAACATCACAATGTAAATATCATTTTTGATGTTT
 CAGCTCACCAGAAGATTCTTACACTTGGGGTAAACACTATCCATGCATTACTTACTGGTAATTACCTGCT
 GGTATATAATTCCATGTAGCCTTTAATATGCTGGGTATCAAATCTGTTCACTGAGTTATGACCAGATA
 AATAATAGATATGCACATGAAAGATGCAAACCTGTGTGATTATTAAAGCCAGCCATGCAGGTCCATGATA
 GAAACAGCAGGTGATGACTCTGCACTCTCATTGTCAAGGTTAGCTATATCCCCAGTTGCAAAACAGCCAG
 ACTTGAGCTGTGCTCTGGTCATCTTTGAGTTTAAAGCCCTTTTGTGTATAAGGCTGTGGAAGTTGTACTC
 CAATGGCTGAAGCCATGTTGTTAATATGGCTGATGGGAGCATCCCTGCAGCTGAACCCAGCACTTTTTAT
 GCTCCCACTGTGGTTGAGCTTTATGTTTACAGTCTCAGCAACAACACTTATGCATCCAAACACTCACAAA
 TGAAACCTGAAAGAATCTTTTCTGAGCCTCTTAAAGAGGAAAATGATGATAACATTAAAGACTCTGAAC
 ACCCAAGGTTGGTGTACATATAAAAAATTAAGCTGATGACTTTGCAGTGACTCAAGTTGTCTCTTTATCA
 TGGTTTACCAGGTAGAGTGCCCTGGCTATTACTATATAATGAAGCCCACTGGCTTGACTTGTAAGTTCAAC
 CTAACCACAATCCTAGACCATCATGGATTTAGGAGTAGATTCTTCTTGAAATCCCACATCCAGAACTA
 GACATTAGAATGTTGAGGCAGTTTCCAGAGAAACAAGCATATTGCCTCATGGATGAAAGACTTGTAAGTT
 CTAGTTTCACTGACTTGTTATATCTACTTACATACAACAGGGAGGCAAGAGGATTCTCTGTCTCTCTGG
 TGACTGAGTGTAATAATATGTGCAAGTCTGCAGCACAGTGACCAATCTGACAATCGAGCTCTGGATCAC
 CACTTGATTATGTAGTAGACTCATTTATAAAGCAGCTTAGGAACTAATTAAACATGGAGGATGAATTACC
 TTCCTATCCCTTGAGATAAGACATCTTTCACTTTTATGATTAAAGGATTGTTGCTGTTTTATAGTTACTCT
 GTTCATCACAGTGTAATGGTGATGCGTGTCGTAGGTGTGCAGCTATTTGAGGGACTAAGGGATGGAGAT
 ATTCTGTCAAATGAATCTCTTCACTATACAGTTTGTGGGAGGGATATGAGACATGTGGATGGCAGTGAG

FIGURE 26.2

63/64

AGATCGTGCCTCTAGATCTTGATGGAGGCTTGGTGAGACACACTTAAATAAGCACGTGGAGGTTAGAATA
GAGGGCAGAGTAAAAGGAAGCTCCATCTGAGCAAGTACACCAAATGATCTCAGCCCTGCAACTTGACCCA
GGTAGGGCCACCACCTACGCCCTTCACTTGTGACCCAAGCTCCAACCACAGAGAGTTTGTGACAAGTTTGTGTT
ATGATGTTGGCTTGGCTTTGTATTTTAAATTAACCTTGGATTTTGTAGTGGTTTGTGCATATAACTGTCCTG
AGTTTGGTAGGTAGGATTACTTTGAAAAGGGTTTACTAGTGTGGTCCTCCGGGTAGAATTTAGCTGTAAC
ATGTTGTTAGCCAGCCTGTAGACTGTTAATTACTTAATAATCTCATTGGGAAAATACTAGTAGTTTTATA
TTTGGATGACATAAATTGGAAAAAGCAGATTAGCTGCTACTACTTTTAAAAGACTTAAGGTCGGGATGCCT
TTTTTCCATGTAAGGAAATGAAAAGACCCAAAATCTTCAGGCAAAAAGCAAGTTGCAAAATTAGAAAACC
ATTGGCTAAAAATGTGTTTTGTGAGTTTCCAAATGGATGAATTTTCATTTGGACATTACATCACTAAAT
TCATTAGATTTTGTCTGCATTGGAAAAGATACTCTTCTAGCATATCTTCCCAAAGATATCTAATTTGGAT
TCTGTTTCATGCAAAATTTGCATCCCGGAGGTTGAAGTTGGAGTTTGGAGTTGGAAAAATATCTTTGAAGGC
AGAATCAGTTGAGTTGTGAGGGTGAAGCCTCACATACTTCTCAACAGACATGATAAAATTCACCTGCATG
AGTTGGCAGGTGGGAGAACCAAAACCTGGATCACTGGGTAAAGACTACTCAGTAAAGCAATGAACCTGCTTGCT
TAGAGAAGCATCACTATCCCCATTGAGAAAAATGTGTGGCAAGATGATACAGCTACACAGTATCAAAATGA
ATGGGTCAATTCAGCACCCCCAAAATTTAATTTCTGTGGGAAAAAATTATTGAGCCAGTTGTCACTGTTCTG
TTACATGACTGGCAGACTAAATTTCTCATCGTTGTTGTTATTGTTGTTGTTGTTTCTCATTTTCACTCGC
ACGGCCTTATTCTCATAATTAATACTAATTCATTTTCTCTTAGTGTAGTAGACTCCAACAACAGAAG
TGGCATCTGTGATTCTAATCAGCATTTACCCTGGCAGGAGACTAATCAGATAGGCCGGTCTCAGACAT
TAATCTTACCATCTGATATTTTGGTGAAGGAAAAAGTATTAATTTCTTTCCATCCTCCTCCTCAGAAA
TATAGAAGCCCTCTTTACCAAAATCATCACATTTTACTCTGTAATCTACCAGCTAAAAGAAAAATTGCATT
GAAGCCCCACAAAGCCAGATTGCAGTTCTTGCCCTTTTTCGCTCTGACATGAGATGTTAAAGAAATTATT
CATTGTGCTCACATTGGGTTAGGGGACACTGAACTGCTTTTGTAGATCCATGATCAGTCATCATCTCTTA
AGAGATTGGAGCTTTGCTGTTTTCATTAACCTGTGTCAGTGTAGACTAATGGTGTTTAATAAAAAATCATTCAA
AATTTCAAACCTCTTTTGCCAGTGACCTCAATTTTGTGCTCTGTGATTTGTATCAGACTTTGAGGAGGG
AAGGGGGAAGTGAAGGAAGCCTACGTCCAGGCCCTGACAGGATGCTGCAGTAGCAAGCTCAAGCTCGCC
TGCTTGCCAGCAGTTGCTGGTGAGCAGCAGCATGCAGACCAGCTGTGGGAAGCCTCCTGAAGAATGCCCC
AGCTGATGCTTTAGCTGGGAATAGTTTGTTCCTATTGGGGAACTCATTGTTCTCCAGTCTCTGCAGCAG
GAAGCCAGCTGTGATATTCGGAGGGAATTTAGATGCTTTACCTTTTGGTTTTGTCTCTGCATCACTCAT
GTGGTACGAAAGTGTCTCTGAGAATAGAGCCCAATGTGGTGACAATGGGTAGTCAAATGCACCCAGAT
GCTCAAGCCCTGTTGTGGTTCTGCAGTGTTTATGAAATTTGGGAGGAAGGAGACCCTGGACATAAGCAAA
ATTGGAGACACTCAACAGAGGCTAAGTTAATGCCGTGTGCGCCAGAACAAGATCTAGCTTCTCATTTGGT
CAGCCTAGCATGCAACCAGTGGTGTGCTGGTAAATGTTTAAACAACCAGCTCGCTGAGAATAGAAAAGCAC
CTGGTTTGCACCATTTGCCAATTTCCATGGCATAAAATACTACCACCTTTAGATGATTTTAAAGCTACCAACT
GTGATGTCACTGAACACATGGTTGGAAAAGAGATGCACGCAGTTGGCTCTTGCAAGCCTGGGCAAAAAATGC
TTCAACACGCCACTGGATGCAGCCAGTCAGAGGGTTCATATTTAATATATGTGTTTATGTGGACACACAC
AGACACACACACACAACTCACCCCTTACACACACACTTCGATGACTAAAACAATTACATAGTTTTAAGAT
ATGAATCAATGTGTGAATGTAGAAAGCTTATGATAAGGCCCTAGAGGTATGGGTGCTTGGAAAGCCTAG
GTTTTAAGCAGGAGAATAGCTGAGAAGAATGAAGCCCTCTGAGCTGAAAGGAGAGATGGATCAATGGAG
ATGGTTCATCATCTCCTTCCATATCTCACAGGTAAATGGGCACTCAGAAAACCCTCACGATTGATTTT
TTAAAAAGATAAGTGAGTGTTTTTTATTTTATTATTATGTCATCATTTTGTATTACAAATGCTATT
TGTAACCTTTACATGTAAGTAAAGTATTTACGGGAACCTATGGAGAATAGACAATCCAGAATT
TACTGTGTTTTTCTTTATGTGACGTGGAACTCAGTAATTTCTCCACCTTCACATTGTTGTTTCTATAAGA
ATTTTACTTTTAGTTATTAGGGAATCTAAGTTTTTTGTTAACATTTGTTTTTGTATTAAAAAGTATCTACTTA
CTGTTTTAGCTCTGAACTCAAACCAGAAATATCTCTGTATCAATTGCATGACTATTACAGAAAACAATAATCC
AAACCAAAATAATTTCTTTTCCACCCAGTACGAAGAAAACTAAGCTCAGTAACAAGAAGGCATAAACTAA
AGTATATAATGAGGCTTTTCATTAAATACACACACACACACACTCACACACACACACATACACTTTTTAAA
TTTTTAAATTAGGCCTCCACACATAAATCATTTTGAAGTAGAATAGAAAAATCTCAAAGAATTCATTCTC
CTGGTCTGTGCATCTTCTGCAGTTAATAAGAGGTTTGTATCTGGAAAGATGGAAGAATTTGTTCTAAAA
TCTTATTTTCAAAAAAAATTTCCATTTTCTCTCTGGGCTGTATCCATGGTTGAATGTTAGCCCTGGA
GGAGATCCATGCTTACTCGCTCTTCTGGCCCTTCTGCTTTTGCCTCTGCAATTTCTTTTGTAGCTGG
CAGGATAGCAGGACTGGGGTCTATCCTTTTCATGGTATTTGCTACAATATTGTCCTTACTGGAAAAATGG
TAACATCCGGGTCTGATTTAATTGGCATTACACTTACACAGGGACTCTGAGCACCCCCGTCAACACACCA

FIGURE 26.3

64/64

GACAGTGGACCAGTTTTTCACAGCTACAAAGAGCTAGAAATGTGTTTAAACATCATCCAGTGCATCCCCTAA
TTCAAAACCATCCTCACTAATCAATCATATTACCCATAAAATATTACAAATGAGATTGATTCCATCTCAA
GACAATTTGTCAAATACTTAATTTTCTTCCTGGATGATTCTACTTACTGGATATTTTAGAAAGAGAAATG
TCTGAGATAAAAATCCCTCACATTTACTCAATATAACAAATTACTGTTTCTACTCCTATTCTGAGTAGTGC
TTCTGAAGATTGTTTGCTGTAGTGTTGTCTTTGATAAAATGAATGTCAGTAGTGAGCCTTTTAGAGATAC
CATGCTCAGACATCCTCTTTGGGATCAGAAGATACCTAAAATTCTCCCCTTTTGCCCACTTGTTAGATG
AGTGATATATTCTTTGGATCCTGCAAAGAAGAGATTGGTTTCTTTTCTTTTCTGGTGGTGGTAGTGGTTG
TATCTGTGGCTGTGATGGTTGTTGTTACTTGTCTCTCTCTCTCTCTGGCTCTGGCTTTTGCTTTCTGCT
AGTGTTCCTTTCTCTTTCCAAACAAATAGTTAAATTAAACGTGAGCTTCTGAATTGTACTTGTTCATACTT
TCAAAACATAACAGATTAATAAAAATAGATGTGTCTGATTTAAACATGCCCCCTGGAAAGGCATGCTG
TATTATGAAATCGTGATAATATAACTGCATTATTACATGGCAGTATAAATATTAGTCTGTTGAATTCATT
TGTCCAATTGTATAACTTTGTGGAGCAGTGTTTTGACCTTTGATACATAATTCTGGAGCAAGTGGAGTGG
TTGCAGGCAGATGAGACAGTGTTATATCAGGATTTTTCAATCAACTTTAGTTGGAGGCCTGGCAATTACA
AACATCTTCAGATGTTTCTGTAACCATTATAAATATGAAAAAACCTCTTCAAAAAATTTCCCATAGTAC
TTCAGTCAAGACTTTTTAGGTTTATCTTTTTTTTTTCAATTTCTCCTTTTCCCTTTCCATTATTTTTTCGAT
GGGGGGTGTGTTATCATTTGACTGAAGAAATATTTTGATTGCAATGGTCTCTCTCTCTCTCCCCCTCTCTC
TCTCTCTCCTCTATTCTTTCCCTCCTTCCCTCTGTCCATCACCCCTCATTAAAAATATTGAAATCTGGAGTC
TTTGATAAATCTGCATTAGACCAGGCTATATGCTAGGAATGAAATCTGGGCAAATATCGATGGGTTTTCA
AAGATGCTCCATGTTTCATTGGGCCCTTTCACACCCACAGTGATAAATGAAAAGGATAGAGGTAGTTTT
TTCAAAAGAGCACTTTAATAATATCCTCTGAGACCTAATGCAGTTTAACAAATGACTCCACCTATTTTTTC
CAGTAGGTAAATTGACTGAGACTTGCAAAATACCCCTGAGAGTTGTGAGGGGTGTCTTCTGCCTGGTCTA
TAGCGTGTGTTTGTGTTGTTATCTAACAGGCACATTACGTCCTCGTGACTCATATGAAGTATTTCCCTA
ACATTCCCATTAGCCTGTATATAAGAATCAGAAAAGATAATCCCAACATGTTGTAAATGAAGATTTGAGCTC
TATAACCTTTCTCTTCTTCTTCTGGAAGAAAAAGGACATTTTTCATGCATATTTTAAACAGAAATTTTGTATA
TTTAAGTGTCTAGAAAATATTTATTGAGTAACGGGACACAAATGGGAATTTAATTGTCTCATATGCT
TTGTGTGTGGGGATGCTTACCAACACCATGTGCTGGACCATTGTGGCAAGCCATAACTGCACAAAGAGT
ACACATCGTCAGTGTGTGTGTGTGTGTGTGTGTGCGCGCACGCACGTGCGTGTGTGTGTCCCTGCATGTG
CAACATGTCTAGCTTGTCTGCTTTCATGGGATTTTAGCTTTCCCTTCTTGAAAAACATTATTTTACAGTT
CCAGGAGGCCCTGGTTACATTACTATATGAAGGCAGTGATTTGAAATGAAATTCCTTTCTCTTGGGAAG
CTTTGGTCATAATATCATGGTTCAATTAAACGGATTCACCCGACTTTGTGATGAAAAAGGCTCTGTAA
AATCCAATTGAGTTTCCAAGAGGAAATTGTAGTAGGTCAAGATGCATGAGAGGGAAGATGGAGGCCACCT
CAGCTGGAGAACATGAGCTGAGTTGAGCCCTCAGTGTTGAAGTTGACTTGCTCCAAGCTGCAGTCTAAAA
CCCTGGGGCCCCGTGCTTGGCCTATGCTCCCTCCCAAGTAAGTAGAGGAGCAGAACCATCAGGAACAGCCT
GCCTGGCTCCTATGAAGAAAACCTTCTGACGTCTGTCCCCAAAGGAAGACCCTTTCCCCAAGGGCACCC
CAGGTGGCCATTAAATTGTGATGATCATTAGAAAAGTGCCCCCTTGGCTTTATGAGAATCCAATTAGTCT
TCTGAACCACCTTTTCTTGGGTGCAGATTTCCAACATTCATGCTCATTGCAGATCCACCAACTGTCACTG
TTCTTAACAAGCATGCTCGTCTTGTGAGAAATTTAGTAAGTTCCAATTTCTGTACAGACCAGGGTAAAC
TGTTCTAAAATCAATCAATTAATGAAATGTTATCTGGTTTTTAAAGCTGGTTTCATGTGCTTTATGTGT
ATAAACTATATCTGCCTGTGTGGCTTTGCATTTCAAATGTGTGGCGCACAAAGCGTTTTGTGGTGCCTT
GTTCTCAGTACAGTAACCTGTGTGACAAACATTTTAAATGTGGTTTTGTGTTTTCCAACAAGATGTCTCT
GTAAAAATGATATTGGCTGAGCTGGTGCGTTGGTTTTCTCTCATAGAGGCATTAACTATACTGCCAATGCA
TTGAATTATTTAAAAATGCAAAATAAAATTTTTATGAAAATCTCA

FIGURE 26.4

LISTE DE SEQUENCES

<110> INSERM
ALLIEL, Patrick
PERIN, Jean-Pierre
RIEGER, Francois

<120> FAMILLE DE SEQUENCES NUCLEIQUES ET DE SEQUENCES
PROTEIQUES DEDUITES PRESENTANT DES MOTIFS RETROVIRAUX
ENDOGENES HUMAINS, LEURS SEQUENCES FLANQUANTES ET LEURS
APPLICATIONS.

<130> 598EXT21

<140>
<141>

<150> 9807920
<151> 1998-06-23

<160> 122

<170> PatentIn Ver. 2.1

<210> 1
<211> 2599
<212> ADN
<213> Homo sapiens

<400> 1
atccccctgcc ttaatcgcca agctccttca ggagaacaaa gaacaggcca ttaccctgga 60
gaagactggc aactgatttt acccacaagc ccaaacctca gggatttcag tatctactag 120
tctgggtaga tacttttcacg gggtgggcag aggccttccc ctgtaggaca gaaaaggccc 180
aagaggtaat aaaggcacta gtccatgaaa taattcccag attcggactt ccccaggcgt 240
tacagagtga caatagccct gctttccagg ccacagtaac ccaggagta tcccaggcgt 300
taggtatacg atatcactta cactgcgcct gaaggccaca gtcctcaggg aaggtcgaga 360
aaatgaatga aacactcaaa ggacatctaa aaaagcaaac ccaggaaacc cacctcacat 420
ggcctgctct gttgcctata gccttaaaaa gaatctgcaa ctttcccaa aaagcaggac 480
ttagcccata cgaaatgctg tatggaaggc ctttcataac caatgacctt gtgcttgacc 540
caagacagcc aacttagttg cagacatcac ctccctagcc aaatatcaac aagttcttaa 600
aacattacaa ggaacctatc cctgagaaga gggaaaagaa ctattccacc cttgtgacat 660
ggtattagtc aagtcacctc cctctaattc cccatcccta gatacatcct gggaaggacc 720
ctaccagtc attttatcta ccccaactgc ggttaaagtg gctggagtgg agtcttgat 780
acatcacact tgagtc aaat cctggatact gccaaaggaa cctgaaaatc caggagacaa 840
cgctagctat tccgtggaac ctctagagga ttgctgcctg ctcttcaaac aacaaccagg 900
aggaaagtaa ctaaaatcat aaatccccat ggccctccct tatcatattt ttctctttac 960
tggtctttta cctcttttca ctctcactgc accccctcca tgccgctgta tgaccagtag 1020
ctccccctac caagagtttc tatggagaat gcagcgctcc ggaaatattg atgccccatc 1080
gtataggagt ctttctaagg gaacccccac ctctactgcc cacaccata tgccccgcaa 1140
ctgctatcac ctgcccactc tttgcatgca tgcaaatact cattattgga caggaaaaat 1200
gattaatcct agttgtcctg gaggacttgg agtcactgtc tgttggaact acttcacca 1260
aactggtatg tctgatggg gtggagtcca agatcaggca agagaaaaac atgtaaaaga 1320
agtaatctcc caactacccc ggttacatgg cactctagc ccctacaaag gactagatct 1380
ctcaaaacta catgaaaccc tccgtaccca tactgcgctg gtaagcctat ttaataccac 1440
cctcactggg ctccatgagg tctcgcccca aaacctact aactgttgga tatgcctccc 1500
cctgaacttc aggccatatg tttcaatccc tgtacctgaa caatggaaca acttcagcac 1560
agaaataaac accacttccg ttttagtagg acctctgtt tccaatctgg aaataacca 1620
tacctcaaac ctcacctgtg taaaatttag caatactaca tacacaacca actcccaatg 1680

```

catcagggtgg gtaactcctc ccacacaaaat agtctgccta ccctcaggaa tatttttttgt 1740
ctgtgggtacc tcagcctatc gttgtttgaa tggctcttca gaatctatgt gcttcctctc 1800
attcttagtg ccccttatga ccatctacac tgaacaagat ttatcacagt atgtcatatc 1860
taagcccccgc aacaaaagag taccatttct tccttttgtt ataggagcag gagtgcctagg 1920
tgcactaggt actggcattg gcggtatcac aacctctact cagttctact acaaaactatc 1980
tcaagaacta aatggggaca tggaacgggt cgccgactcc ctgggtcacct tgcaagatca 2040
acttaactcc ctagcagcag tagtccttca aaatcgaaga gcttttagact tgctaaccgc 2100
tgaaagaggg ggaacctgtt tatttttagg ggaagaatgc tgttattatg ttaatcaatc 2160
cggaatcgtc actgagaaag ttaaagaaat tcgagatcga atacaacgta gaggagagga 2220
gcttcgaaac actggaccct ggggcctcct cagccaatgg atgccctgga ttctcccctt 2280
cttaggacct ctagcagcta taatattgct actcctcttt ggaccctgta tctttaacct 2340
ccttggttaac tttgtctctt ccagaatcga agctgtaaaa ctacaaatgg agcccaagat 2400
gcagtccaag actaagatct accgcagacc cctggaccgg cctgctagcc cactgatctga 2460
tggttaatga atcaaaggca cccctcctga ggaaatctca gctgcacaa cctctactacg 2520
cccaattga gcaggaagca gttagagcgg tctcggccaa cctccccaac agcacttagg 2580
ttttcctgtt gagatgggg 2599

```

<210> 2

<211> 1326

<212> ADN

<213> Homo sapiens

<400> 2

```

gccgcctggc actcctgagg gaagtataaa ttataacacc atcttacagc tagacctctt 60
ttgtagaaaa ggcaaatgga gtgaagtgcc ataagtacaa actttctttt cattaagaga 120
caactcaciaa ttatgtaaaa agtgtgattt atgccctaca ggaagccttc agagtctacc 180
tccctatccc agcatccccg actccttccc caactaataa ggacccccct tcaacccaaa 240
tgggtccaaa ggagatagac aaaagggtaa acagtgaacc aaagagtgcc aatattcccc 300
aattatgacc cctccaagca gtgggaggaa gagaattcgg ccagccaga gtgcatgtgc 360
ctttttctct cccagactta aagcaaataa aaacagactt aggtaaattc tcagataacc 420
ctgatggcta tattgatgtt ttacaagggt taggacaatt ctttgatctg acatggagag 480
atataatgtc actgctaaat cagacactaa ccccaaatga gagaagtgcc accataactg 540
cagcctgaga gtttggcgat ctctggtatc tcagtcaggc caatgatagg atgacaacag 600
aggaaagaga atgattcccc acaggccagc aggcagttcc cagtctagac cctcattggg 660
acacagaatc agaacatgga gattggtgct gcagacattt gctaacttgt gtgctagaag 720
gactaaggaa aactaggaag aagtctatga attactcaat gatgtccacc ataacacagg 780
gaagggaaga aaatcctact gcctttctgg agagactaag ggaggcattg aggaagcgtg 840
cctctctgtc acctgactct tctgaaggcc aactaatctt aaagcgtaag tttatcactc 900
agtcagctgc agacattaga aaaaaacttc aaaagtctgc cgtaggcccg gagcaaaact 960
tagaaacctt attgaacttg gcaacctcgg ttttttataa tagagatcag gaggagcagg 1020
cggaacagga caaacgggat taaaaaaaag gccaccgctt tagtcatgac cctcaggcaa 1080
gtggactttg gaggtcttgg aaaagggaag agctgggcaa attgaatgcc taatagggct 1140
tgcttccagt gcggtctaca aggacacttt aaaaaagatt gtccaagtag aagtaagccg 1200
ccccctcgtc catgcccctt atttcaaggg aatcactgga aggccactg cccaggggga 1260
caaaggctct ctgagtcaga agccactaac cagatgatcc agcagcagga ctgagggtgc 1320
ctgggg 1326

```

<210> 3

<211> 10499

<212> ADN

<213> Homo sapiens

<400> 3

```

ccctggggcg ggcttccttt ctgggatgag ggcaaaacgc ctggagatac agcaattatc 60
ttgcaactga gagacaggac tagctggatt tcctaggccg actaagaatc cctaagccta 120
gctgggaagg tgaccacgtc cacctttaaa cacggggcct gcaacttagc tcacacctga 180

```

ccaatcagag	agctcactaa	aatgctaatt	aggcaaagac	aggaggtaaa	gaaatagcca	240
atcatctatt	gcctgagagc	acagcaggag	ggacaacaat	cgggatataa	acccaggcat	300
tcgagctggc	aacagcagcc	cccctttggg	tcccttcctt	ttgtatggga	gctgttttca	360
tgctatttca	ctctattaaa	tcttgcaact	gcactcttct	gggccatgtt	tcttacgggt	420
cgagctgagc	ttttgtcac	cgccaccac	tgctgtttgc	caccaccgca	gacctgccgc	480
tgactcccat	ccctctggat	cctgcagggt	gtccgctgtg	ctcctgatcc	agcgaggcgc	540
ccattgccgc	tcccaattgg	gctaaaggct	tgccattgtt	cctgcacggc	taagtgcctg	600
ggtttgttct	aattgagctg	aacacttagt	actgggttcc	atggttctct	tctgtgacct	660
acggcttcta	atagaactat	aacacttacc	acatggccca	agattccatt	ccttgaatc	720
cgtgaggcca	agaactccag	gtcagagaat	acgaggcttg	ccaccatctt	ggaagcggcc	780
tgctaccatc	ttggaagtgg	ttcaccacca	tcttgggagc	tctgtgagca	aggaccccc	840
ggtaacattt	tggcaaccac	gaacggacat	ccaaagtgtg	gagtaatatt	ggaccacttt	900
cacttgctat	tctgtcctat	ccttccttag	aattggagga	aaataccggg	cacttgctcg	960
ccagttaaaa	acgattagt	tggccaccgg	acttaagact	caggtgtgag	gctatctggg	1020
gaagggtctt	ctaacaacct	ccaacccttc	tgggttgggg	acttggtttg	cctcaagcca	1080
gcttccactt	tcagttttct	tggggaagcc	gaggcccgac	tagaggcgaga	aagctgtcgt	1140
cctgaactcc	cggcagtagc	cggttgagat	catggtgtag	ccagaagtct	caacagtcgc	1200
ccatgcatgc	acccctatct	ttccttctga	cccatacctc	ctgggtccca	accacaactt	1260
tcttcaaaagt	gtagccccaa	aattctcctt	acctctgaat	atacttcctc	tgatccctgc	1320
ctcctaggtg	ctattgggtc	agacttccat	ttcctctagc	aagttgtatc	tccaaaggga	1380
tctaaggaa	ctctgcgctg	cgctccttag	cacttaggct	ataacccagg	gagtcttatc	1440
cctggtgtcc	ctcccaattt	aggcatacag	ctcttgacat	gggcagttat	gtaggaccca	1500
ctccccacca	cccttgccag	ggccccaagt	ttgtaaatgg	ctgagggaaa	agagagacag	1560
aggagagaga	gagaaatgga	ggagaaagag	agagagacag	agaggagaga	gagacagtga	1620
gagagacaga	agagagagag	agacaaagag	gagagagaga	gagtaaaaga	gagaaagaaa	1680
gagaaagaaa	tagtaaaaaa	cagtgtgccc	tattccttta	aaagccaggg	taaattttaaa	1740
acctgtactt	gataattgaa	ggtcttctct	gtgacctat	agcactccaa	tccactttgt	1800
ggtcagtgtg	aataagagca	taggccgaaa	gcactgaggc	cattgacaac	ccgtagcttc	1860
cctatcaaaa	atccttaacc	cagtaacccg	cagatggacc	aaatgcattc	agtcggtagc	1920
gcaactgctt	tgctaaaaagt	agaaaaagta	cttttagagg	aaacctcatt	gtgagcacac	1980
ctcacctgtt	cagaattatt	ctaataaaaa	aagcaaaaag	gtagcttact	aactcaaaaa	2040
tcttaaaagta	tggggctatt	ctgttagaaa	aaggtaatgt	aactccaacc	actgataatt	2100
cccttaaccc	agcagatttc	ctaaccggat	ttaaatctta	attaccatac	aaaggtccga	2160
ccagacctag	gcggaactcc	cttcaggaca	ggacgataga	tggttcctcc	cagggtgattg	2220
aggaaaaaaa	ccacaatggg	tattcagtaa	ttgatacggg	gactcttgtg	gaagcagagt	2280
tagaaaaatt	gcctaataac	tgggtctcctc	aaacgtgtga	gctgtttgca	ctcagccaa	2340
ccttaaaagta	cttacagaat	caaaaagacta	tctcaatcct	gattcaaaa	gttagctaca	2400
ccctctctgt	aatgcatttg	cataagaact	tgtttatggg	aatgcattct	gatggggcag	2460
ctgggttgtt	ataaaaatagg	aaccagacc	agctctagga	ctcacccttg	agcgcaaacg	2520
caatggttgg	catgctggtg	aaggaccact	agaatccagc	agccagacc	cctttctttg	2580
tgggtcaagaa	aggcgggaaa	aggggtgcag	gactgctaca	tcggtaaagca	taactaatcc	2640
gataaacaga	ggtccatggg	tggttacgca	ccctggaaa	gaactcacc	ctgagcaca	2700
aggcaatgtt	gggcacgctg	gtaaaaggacc	actagaatcc	agcagcctgg	acccctttct	2760
ttgtggtcaa	gagaggcagg	aaaacagggtg	caggactgca	acatcagtga	gcataactaa	2820
ttcgataaagc	agaggtccat	gggtggtgat	gcaccctgga	aagaataaagc	attaggacca	2880
tagaggacac	tcagggacta	aagctcatcg	gaaaatgact	aggggttgcg	gcatccctat	2940
gttctttttt	cagatgggaa	acgttccccg	caagacaaaa	acgcccctaa	gacgtattct	3000
ggagaatttg	gaccaatttg	accctcagac	actaagaaa	aaacgactta	tattcttctg	3060
cagtgccgcc	tggcactcct	gagggaaagta	taaattataa	caccatctta	cagctagacc	3120
tctttttag	aaaaggcaaa	tggagtgaag	tgccataagt	acaaactttc	ttttcattaa	3180
gagacaactc	acaattatgt	aaaaagtgtg	atttatgccc	tacaggaagc	cttcagagtc	3240
tacctcccta	tcccagcatc	cccgactcct	tccccaaacta	ataaggacc	cccttcaacc	3300
caaatggtcc	aaaaggagat	agacaaaagg	gtaaacagtg	aaccaaagag	tgccaatatt	3360
ccccaatat	gaccctcca	agcagtggga	ggaagagaat	tcggcccagc	cagagtgcac	3420
gtgccttttt	ctctcccaga	cttaaaagcaa	ataaaaacag	acttaggtaa	attctcagat	3480
aaccctgatg	gctatattga	tgttttacaa	gggttaggac	aattctttga	tctgacattg	3540
agagatataa	tgtcactgct	aaatcagaca	ctaaccctaa	atgagagaag	tgccaccata	3600
actgcagcct	gagagtgttg	cgatctctgg	tatctcagtc	aggtcaatga	taggatgaca	3660

acagaggaaa	gagaatgatt	ccccacaggg	cagcagggcag	ttcccagttc	agaccctcat	3720
tgggacacag	aatcagaaca	tggagattgg	tgctgcagac	atttgctaac	ttgtgtgcta	3780
gaaggactaa	ggaaaaactag	gaagaagtct	atgaattact	caatgatgtc	caccataaca	3840
cagggaaggg	aagaaaatcc	tactgccttt	ctggagagac	taaggaggag	attgagggaag	3900
cgtgcctctc	tgtcacctga	ctcttctgaa	ggccaaactaa	tcttaaagcg	taagttttatc	3960
actcagtcag	ctgcagacat	tagaaaaaaa	cttcaaaagt	ctgccgtagg	cccgagagcaa	4020
aacttagaaa	ccctattgaa	cttggcaacc	tccgtttttt	ataatagaga	tcaggaggag	4080
caggcggaac	aggacaaacg	ggattaaaaa	aaaggccacc	gcttttagtca	tgaccctcag	4140
gcaagtggac	tttggaggct	ctggaaaagg	gaaaagctgg	gcaaattgaa	tgctaataag	4200
ggcttgcttc	cagtgcggtc	tacaaggaca	ctttaaaaaa	gattgtccaa	gtagaagtaa	4260
gcgcgcccc	cgtccatgcc	ccttattttca	agggaatcac	tggaaggccc	actgccccag	4320
gggacaaaag	tcctctgagt	cagaagccac	taaccagatg	atccagcagc	aggactgagg	4380
gtgcctgggg	caagcgccat	cccatgccat	caccctcaca	gagccctggg	tatgcttgac	4440
cattgagggg	caggagggtg	tctcctggac	actggtgcgg	tcttcttagt	cttactcttc	4500
tgccccggac	aactgtctc	cagatctgtc	actatctgag	ggggtcctaa	gacgggcagt	4560
cactagatac	ttctcccagc	cactaagtta	tgactgggga	gctttattct	tttcacatgc	4620
ttttctaatt	atgcttgaaa	gccccactac	cttggttaggg	agagacattc	tagcaaaagc	4680
agggggccatt	atacacctga	acataggaga	aggaaacccc	gtttgttgtc	ccctgcttga	4740
ggaaggaatt	aatcctgaag	tctgggcaac	agaaggacaa	tatggacgag	caaagaatgc	4800
ccgtcctggt	caagttaaac	taaaggattc	cacctccttt	ccctaccaaa	ggcagttacc	4860
cctcagaccc	aaggcccaac	aaggactcca	aaagattggt	aaggacctaa	aagcccaagg	4920
cctagtaaaa	ccatgcagta	accctgcag	tactccaatt	ttaggagtag	agaaacccaa	4980
cagacagtgg	aggttagtgc	aagatctcag	gattatcaat	gaggctgttg	ttcctctata	5040
gccagctgta	cctagccctt	atactctgct	ttcccaaata	ccagaggaag	cagagtgggt	5100
tacagtcctg	gaccttcagg	atgccttctt	ctgcacccct	gtacatcctg	actctcaatt	5160
cttgtttgcc	tttgaagata	cttcaaaccc	aacatctcaa	ctcacctgga	ctattttacc	5220
ccaagggttc	agggatagtc	cccatctatt	tggccaggca	ttagcccaag	acttgagcca	5280
atcctcatat	ctggacactt	gtccttcggg	agggtgatga	tttacttttg	gccgcccatt	5340
cagaaacctt	gtgccatcaa	gccacccaag	cgctcttcaa	tttctctgct	acctgtgggt	5400
acatggtttc	caaaccaaaag	gtcacaactc	gtcacagca	ggttacttag	ggctaaaatt	5460
atccaaaggc	accaggggccc	tcagtggaga	acacatccag	cctatactgg	cttatectca	5520
tccaaaaacc	ctaaagcaac	taaggggatt	ccttggcgta	ataggtttct	gccgaaaatg	5580
gattcccagg	tatggcgaaa	tagccaggtc	attaaataca	ctaattaagg	aaactcagaa	5640
agccaatacc	catttagtaa	gatggacaac	tgaagtagaa	gtggctttcc	aggccctaac	5700
ccaagcccca	gtgttaagtt	tgccaacagg	gcaagacttt	tcttcataat	tcacagaaaa	5760
aacaggaata	gctctaggag	tccttacaca	gatccgaggg	atgagcttgc	aacctgtggc	5820
atacctgact	aaggaaaattg	atgtagtggc	aaagggttga	cctcattgtt	tacgggtagt	5880
ggtggcagta	gcagtcttag	tatctgaagc	agttaaaata	atacagggaa	gagatcttac	5940
tgtgtggaca	tctcatgatg	tgaatggcat	actcactgct	aaaggagact	tgtggctgtc	6000
agacaactgt	ttactttaat	gtcaggctct	attacttgaa	gggccagtgc	tgcgactgtg	6060
cacttgtgca	actcttaacc	cagccacatt	tcttccagac	aatgaagaaa	agataaaaca	6120
taactgtcaa	caagtaattt	ctcaaaccta	tgccactcga	ggggaccttt	tagaggttcc	6180
tttgactgat	cccgaacctca	acttgtatata	tgatggaagt	tcctttgtag	aaaaaggact	6240
tcgaaaagtg	gggtatgcag	tggtcagtga	taatggaata	cttgaaagta	atccctcac	6300
tccaggaact	agtgctcagc	tagcagaact	aatagccctc	acttgggcac	tagaattagg	6360
agaagaaaaa	agggcaaata	tatatacaga	ctctaaatat	gcttacctag	tcctccatgc	6420
ccatgcagca	atatggaaaag	aaagggaatt	cctaacttct	gagagaacac	ctatcaaaaa	6480
tcaggaagcc	attaggaaat	tattattggc	tgtacagaaa	cctaaagagg	tggcagtcct	6540
acactgccgg	ggtcatcaga	aaggaaaagga	aagggaataa	gaagagaact	gccaagcaga	6600
tattgaagcc	aaaagagctg	caaggcagga	ccctccatta	gaaatgctta	taaaacaacc	6660
cctagtatag	ggtaatcccc	tccgggaaac	caagccccag	tactcagcag	gagaaacaga	6720
atggggaaac	tcacgaggac	agttttctcc	cctcgggacg	gctagccact	gaagaaggga	6780
aaatactttt	gcctgcaact	atccaatgga	aattacttaa	aacccttcat	caaacctttc	6840
acttaggcac	cgatagcacc	catcagatgg	ccaaatcatt	atttactgga	ccaggccttt	6900
tcaaaactat	caagcagata	gtcagggcct	gtgaagtgtg	ccagagaaat	aatcccttgc	6960
cttatcgcca	agctccttca	ggagaacaaa	gaacaggcca	ttacctgga	gaagactggc	7020
aactgatttt	accacaagc	ccaaacctca	gggatttcag	tatctactag	tctgggtaga	7080
tactttcacg	ggttgggcag	aggccttccc	ctgtaggaca	gaaaaggccc	aagaggtaat	7140

aaaggcacta	gttcatgaaa	taattccag	attcggactt	ccccgaggt	tacagagtga	7200
caatagccct	gctttccagg	ccacagtaac	ccaggagta	tcccaggcgt	taggtatacg	7260
atatcactta	cactgcgcct	gaaggccaca	gtcctcaggg	aaggtcgaga	aaatgaatga	7320
aacactcaaa	ggacatctaa	aaaagcaaac	ccaggaaacc	cacctcacat	ggcctgctct	7380
gttgccctata	gccttaaaaa	gaatctgcaa	ctttccccaa	aaagcaggac	ttagcccata	7440
cgaaatgctg	tatggaaggc	ccttcataac	caatgacctt	gtgcttgacc	caagacagcc	7500
aacttagttg	cagacatcac	ctccttagcc	aaatatcaac	aagttcttaa	aacattacaa	7560
ggaacctatc	cctgagaaga	gggaaaagaa	ctattccacc	cttgtgacat	ggtattagtc	7620
aagtcccttc	cctctaattc	cccatcccta	gatacatcct	gggaaggacc	ctaccagtc	7680
atthttatcta	ccccaaactgc	ggttaaagtg	gctggagtgg	agtcttggat	acatcacact	7740
tgagtcaaat	cctggatact	gccaaaggaa	cctgaaaatc	caggagacaa	cgctagctat	7800
tcctgtgaac	ctctagagga	tttgcgccctg	ctcttcaaac	aacaaccagg	aggaaagtaa	7860
ctaaaatcat	aaatccccat	ggccctccct	tatcatatth	ttctctttac	tgttctttta	7920
ccctctttca	ctctcactgc	acccctccca	tgcgcgtgta	tgaccagtgc	ctccccttac	7980
caagagtttc	tatggagaat	gcagcgtccc	ggaaatattg	atgccccatc	gtataggagt	8040
ctttctaaagg	gaacccccac	cttcaactgcc	cacacccata	tgccccgcaa	ctgctatcac	8100
tctgccactc	tttgcatgca	tgcaaatact	cattattgga	caggaaaaat	gattaatcct	8160
agttgtcctg	gaggacttgg	agtcactgtc	tgttggaactt	acttcaccca	aactgggatg	8220
tctgatgggg	gtggagtcca	agatcaggga	agagaaaaac	atgtaaaaaga	agtaatctcc	8280
caactcacc	gggtacatgg	cacctctagc	ccctacaaag	gactagatct	ctcaaaacta	8340
catgaaaccc	tccgtaccca	tactcgccctg	gtaagcctat	ttaataccac	cctcactggg	8400
ctccatgagg	tctcggccca	aaaccctact	aactgttggg	tatgcctccc	cctgaacttc	8460
aggccatatg	tttcaatccc	tgtacctgaa	caatggaaca	acttcagcac	agaaataaac	8520
accacttccg	ttttagtagg	acctcttggt	tccaatctgg	aaataaccca	tacctcaaac	8580
ctcacctgtg	taaaatttag	caatactaca	tacacaacca	actcccaatg	catcaggtgg	8640
gtaactcctc	ccacacaaat	agtctgccta	ccctcaggaa	tattttttgt	ctgtgggtacc	8700
tcagcctatc	gttggtttgaa	tggtctcttca	gaatctatgt	gcttcctctc	attcttagtg	8760
ccccctatga	ccatctacac	tgaacaagat	ttatacagtt	atgtcatatc	taagccccgc	8820
aacaaaagag	taccattctc	tcctttttgtt	ataggagcag	gagtgcctag	tgactaggt	8880
actggcattg	gcggtatcac	aacctctact	cagttctact	acaaactatc	tcaagaacta	8940
aatggggaca	tggaacgggt	cgccgactcc	ctggctacct	tgcaagatca	acttaactcc	9000
ctagcagcag	tagtccctca	aaatcgaaga	gcttttagact	tgctaaccgc	tgaaagaggg	9060
ggaacctgtt	tattttttagg	ggaagaatgc	tgttattatg	ttaatcaatc	cggaatcgtc	9120
actgagaaag	ttaaagaaat	tcgagatcga	atacaacgta	gagcagagga	gcttcgaaac	9180
actggaccct	ggggcctcct	cagccaatgg	atgccctgga	ttctccccctt	cttaggacct	9240
ctagcagcta	taatattgct	actcctcttt	ggaccctgta	tcttttaacct	ccttggttaac	9300
tttgtctctt	ccagaatcga	agctgtaaaa	ctacaaatgg	agcccaagat	gcagtccaag	9360
actaagatct	accgcagacc	cctggaccgg	cctctagctg	cacgatctga	tgttaatgac	9420
atcaaaggca	ccccctcctga	ggaaatctca	gctgcacaa	ctctactacg	ccccaaattca	9480
gcaggaagca	gttagagcgg	tctcggccaa	cctccccaac	agcacttagg	ttttcctggt	9540
gagatggggg	actgagagac	aggactagct	ggatttccta	ggctgactaa	gaatccctaa	9600
gcctagctgg	gaaggtgacc	acatccacct	ttaaacacgg	ggcttgcaac	ttagctcaca	9660
cctgaccaat	cagagagctc	actaaaatgc	taatttaggca	aagacaggag	gtaaagaaat	9720
agccaatcat	ctattgcctg	agagcacagc	aggagggaca	atgatcgga	tataaaacca	9780
agtcttcgag	ccggcaacgg	caacccccctt	tgggtccctt	ccctttgtat	gggagctctg	9840
ttttcatgct	atttcaactct	attaaatctt	gcaactgcac	tcttctgggtc	catgtttctt	9900
acggcttgag	ctgagctttc	gctcgccatc	caccactgct	gtttgccgcc	accgcagacc	9960
cgccgctgac	tcccatccct	ctggatcatg	cagggtgtcc	gctgtgctcc	tgatccagcg	10020
aggcacccat	tgccgctccc	aatcgggcta	aaggcttgcc	attgttctctg	catggctaag	10080
tgccctgggtt	catectaatt	gagctgaaca	ctagtcactg	ggttccatgg	ttctcttctg	10140
tgaccacag	cttctaatag	agctataaca	ctcacgcgat	ggcccaagggt	tccattcctt	10200
gaatccataa	ggccaagaac	cccaggtcag	agaacacgag	gcttgccacc	atcttgggag	10260
ctctgtgagc	aaggaccccc	aagtaacaca	accatgaggg	tgcaaatagca	tgggcccacta	10320
atggtagagc	aagaaaacag	aagggccctg	gttcctcgaa	ggcatcagtg	agctgaaatg	10380
cctgccttgg	atgtcctatt	cctaggtgtt	tttctgcctg	aagcagatta	aaccttttgt	10440
tcacttctcc	aagtagggct	tctattacag	cccaaatcaa	tccccacccc	agatgacat	10499

<210> 4
 <211> 2784
 <212> ADN
 <213> Homo sapiens

<400> 4
 ctcccttcagg agaacaaaga acaggccact acccaagaga agactggcaa ctagatttta 60
 cccatatgcc caaatctcag ggatttcagt atctactagt ttgggtagat actttcactg 120
 gttgggcaga ggccttcccc tgtaggacag aaaaggccca agaggtaata aacgttcagt 180
 aaataattcc cagattcgga cttccccaag gcttacagag tgacaatggc cctgctttca 240
 aggctacagt aaccaagga gtatcccagg tgttagggtat acaatatcac tcacactgag 300
 cctggaggcc acagtcctca ggaaagggtg agaaaatgaa caaaacactc aaatgacatc 360
 taaaaaagct aatccaggaa acccacctcg catggcctgc tctgttgccct atagccttac 420
 taagaatccg aaactctccc caaaaagcag gacttagtcc atacaaaatg ctgtatggac 480
 ggcccttcct aaccaatgaa cttgggcttg accgagagac agccaactta gttgcagaca 540
 tcatctcctt agccaaatat caacagggttc ttaaaacatt acaggggagcc tgtccccaag 600
 aagagggaaa ggaactattc caccctggtg acatggtatt agtcaagtcc cttccctcta 660
 attccccatc cctagatata tcttggaag gaaactaccc agccatttta tctaccctaa 720
 cggcagttaa agtggtctga gcggagctct ggatacatca cactcaagtc aaacctgga 780
 tactgccaaa ggaactcaaa aatccatgag acaatgctag ctattcctgt gaacctctag 840
 aggatctgag cctgctcttc aaatgacaac cagggggaaa gtaactaaaa tctgtaaatcc 900
 cctggccctc ccttatcata tttttctctt tactgttctc ttacccctt tcaactctac 960
 tgcaccccggt ccatgccact gcaccccggt catgcccgt ccatgccagt agctccctt 1020
 agcaagagtt tctatggaga atgcagcgtc ccggaaatat tgatgcccc ttgtatagga 1080
 gtttatctaa gggaaccccc accttcaactg cccacaccca tatgccccac aactgctata 1140
 actctgccac tctttgcatg catgcaaata ctctatttg gacaggaaaa acgattaatc 1200
 ccagttgtcc tggaggactt ggaggactca cttcactcat accagtatgt ctgatggggg 1260
 tggagttcaa gatcaggcaa cagaaaaaca cataaaggaa gtaatctccc aactgacctg 1320
 ggtacatagc acccctggcc cctacaaagg actagatctc tcaaaactac atgaaacct 1380
 ccat acccat actggcctgg taagcctatt taataccacc ctgactgggc tccatgaggt 1440
 ctgggccccaa aacctacta actgttggat gtgcctcccc ctgcacttta ggccatacat 1500
 ttcaatccct atacctgaac aatggaacaa cttcagcaca gaaataaaca ccacttctgt 1560
 tttagtaggt cctctttcca atctggaat aaccatacc tcaaacctca cctgtgtaaa 1620
 atttagcaat actatagaca cagccaactc ccaatgcac aggtgggtaa ctccctccac 1680
 acgaatagtc tgcctaccct caggaatatt ttttgtctgt ggtacctcag cctatcattg 1740
 tttgaatggc tcttcagaat ctgtgtgctt cctctcattc ttagtggccc ctatgcccat 1800
 ctacactgaa caagatttat acaatcatgt catacctaag ccccgcaaca aaagagtacc 1860
 cattcttctt tttgttattg gagcaggagt gctaggcgga gtagctactg gcattggcgg 1920
 tatcacaacc tctactcagt tctactacaa actgtctcaa gaactaaatg gtgacatgga 1980
 atgggtcgct gataccctgg tcaccttgca agatcaactt aactccctag cagcagtagt 2040
 ccttcaaaat cgaagagctt tagacttgct aaccgcggaa agcgggggaa cctttttatt 2100
 tttagaggaa aaatgctgtt gttatgttaa tcaatccgga atcatcacc agaaagttaa 2160
 agaaattcaa ggtcgaatat aacgtagagc aaaggagctg caaaacactg gaccctgggg 2220
 cctcctcagc caatggatgc cctggattct ccccttctta ggacctctag cagctataat 2280
 attgttactc ctctttggac cctgtatctt taacctcctt gtttaagtttg tcttttccag 2340
 aatcgaagca gtaaaactac aaatcgcttct tcaaatggag cccagatgc agtccatgag 2400
 taaaatctac cacggacccc tggacgggcc tgctagccca tgctctgatg ttaatgacat 2460
 caaaggcacc cctcccgagg aaatctcaac tgcacaaact ctactacgcc ccaattcagc 2520
 aggaagcagt tagagtgggt gttggccaac ctccccacaa gcagttgggt tttcctggtg 2580
 agagggggga ctgagagaca ggaataacta gatttcttag accaactaag aatccctaag 2640
 actagctggg aaggtgaccg cttccacctt taaccacgg gcttgcaact tatctcacgc 2700
 ccaaccaatc agatactaaa gagagctcac taaaatgcta attaggcaaa aacaggagat 2760
 aaagaaatag ccaatcatct gttg 2784

<210> 5
 <211> 1799
 <212> ADN
 <213> Homo sapiens

<400> 5
 gggattctta gtcggcctag gaaatccagc taatcctgtc tctcagtcctc cccactcaac 60
 aggaaaaccc aagtgcctgtt ggggaggttg gctgacgacc agtctaactg ctctcctgcgg 120
 aattggggca tagtaggggt tgtgcagttg agatttcctc gggaggggtg cgttcgata 180
 cattacaatt ggagcatggg ctagtaggcc ggtccagggg tccacggtag atcttagtca 240
 tggacttcat ctgggggttcc atttgaagaa cgatttgtag ctttacaact ttgattctgg 300
 aagagacaaa cttaacaagg aggttaaaga tacaggggtc aaagaggagt atcaatatta 360
 gagctgctag agatcctaag aaggggagaa tccagggcat ccattggctg aggaggcccc 420
 aggggtctggt gtttttgaag ctctctgtt ctacgttgta ttcaatctcg aatttcttca 480
 actttctctg tgacaattca ggattgatta acataataac aacattcttc cgctaaaata 540
 acataataac aacattcttc ccctaaaaat aaacagcttc ccctctttc agagggttagc 600
 aagtctaaag ctcttcaatt ttgaaggact actgatgcta ggaagttaag ttgatcttgc 660
 aagggtgacca gggagtcggc aacccattcc atgtcaccat tgagttcttg agatagtttg 720
 tagtagaact gagtagaggt tgtggtaccg ccaatgccag aacctagtcc acctagcact 780
 cctgctccga taacaaaagg aagaatgagt actcttttgt tgtggggctt aggtacaaca 840
 taattgtata aatcttgttc agtgtaaatg gtcatggggg cactaagaat gagaggaagc 900
 acatagattc tgaagagcca ttcaaacaac gataggctaa ggtaccacag acaaaaaata 960
 ttctgagggg taggcagact attcgtgtgg gaggagttac ccacctgatg cattgggagt 1020
 tggttgtgtc tacagtattg cttaaatttta cacaggtgag gtttgaggta tgggttattt 1080
 ccagattgga aacaagaggt cctactaaaa cggaagtggg gtttatttct gtgctgtagt 1140
 tgttccattg ttcagggtaca gggattgaaa tgcattggcct gaaatacagg gggaggcaca 1200
 accaacagtt agtaggggtt tggaccgaga cctcatggag ccagtgagg gtggtattaa 1260
 ataggcttac caggcaagta tgggtatgga ggggttcatg tagttttaag agatctagtc 1320
 ctttgtaggg gctaggggtg ctatgtaccc gggtcagttg ggaggttact tctttacat 1380
 gtttttctct tgctgatct tgaactccac cccctcaga cataccagta tgggtgaagt 1440
 aagtccgaca gacagtggct ccaagtcttc caggacaact aggattaatc attttccctg 1500
 tccaataatg agtatttgca tgcattgcaa gagtggcaga gttatagcag ttgtggggca 1560
 tatgggtgtg ggcagtgaag gtggagtttc ctttaggtaa actcctattt gatggggcat 1620
 caatatttct gggaagccgc attcttcata gaaactcttg gtaaggggag ctgctgggtg 1680
 tacagcagca tggaggggtt gcagtgagag tgaaaggggg taagagaaca gtaagagaa 1740
 aaatatgata agggagggcc atggggattt acgattttag ttactttcct cacggttgt 1799

<210> 6
 <211> 1489
 <212> ADN
 <213> Homo sapiens

<400> 6
 tgggtgcttg cccgggcact ctgagtcctg ctgctggatc atctgggttag tggcttctga 60
 ctcagaggac ctacgtcccc tggggcagtg ggccttacag tgattccctt gacacgaggt 120
 gcatggacga gggggcggct tatttctatt tggacaatct tttttaaagt gtcctttagt 180
 accgcactgg aagcaaacc cttataggcat ttgatttgcc tagcttttcc cttttccagt 240
 gcctccaaag tccgcttgcc tgagggccat gactaaagcg gtggcctttt ttttatccca 300
 tttgtcccat tctgcctgct catcctgatc tctattataa aaaactgagg ttgccaaagt 360
 caataggggt tctaagtttt gttccgggac taaggcagac ttttgaagtt ttttccta 420
 gctgtagct gactgagtga taaacttatc ctttaagatt agttggcctt cagttagatc 480
 agttgacaga gagaggtatg cttcctcaat gcctccgtta gtactccag aaaggcggta 540
 ggattttctt cttttccctg tgttatagtg gacatcattg aataactcac aggcttctt 600
 ctagttttcc ttagtccttc tagcacgcaa gtttagcaat gtctgcggca ccaatctcca 660
 tgttctgatt ctgtgtccca gtgaggtct acactgggaa ctgctgctg gcctgtgggg 720
 aatcgttctt tttcctctgt tgtcgaccta tcattgacct gactgagata ccagagatcg 780
 ccaaactctc aggcgtgcagt tacggcgaca cttctgtcat ttgggggttag tgtctgattt 840

agcagtaaca	ttatatctct	ccatatcaga	tcaaaggatt	gtcctaaacc	ttgtaaaaca	900
tcaatatagc	cattaggggt	atctgagaat	ttacctaggt	ctattttaat	ttaaagtctg	960
ggagagaaaa	aggcacatgc	actctggctg	ggccgaattc	tcttccctcc	actgcgtctg	1020
agagagaaaa	aggtagctgc	actctggctg	ggccgaattc	tcttccctcc	gcttggaggg	1080
ggcataatcg	gggaatattg	gcattctttg	gttagttggt	tacctctttg	tctatctcct	1140
tttggaccgt	ttgggttgaa	ggggggctct	tattatttgg	ggaaggagtc	tgggggatgc	1200
tggggtaggg	aggtagactc	tgagggtctc	ctgtagggca	taaatcacac	tttttacata	1260
attgcgagtt	gtctcttaat	gaaaagaaa	tttgtagcta	tgacacttca	caccatttgc	1320
cttcttttct	acaaaagagg	tctagctgta	agatggtgtt	ataatttatg	cttccctcag	1380
gatgccaggt	ttctccccct	taaagagtat	atcgttgcca	ggcgttactg	cagaagaata	1440
tgtctttttt	ttcttagcat	ctgagagtca	aattgggtcc	aattctcca		1489

<210> 7

<211> 1216

<212> ADN

<213> Homo sapiens

<400> 7

taaagataca	gggattgaaa	tgtatggcct	gaagtgcagg	gtcatatagg	tgtgggtggg	60
gaaaatgggg	tttcctttag	aaaaactcct	atacgatggg	tcatcaatat	ttccaggaag	120
ccgcattctc	catagaagct	cttggtaatg	ggagctactg	gtagtacagt	ggcatggagg	180
gggtgcagtg	agagtgaaa	agggtaaaa	aacagtaaa	agaaaaatat	gataaggagg	240
gggttcagtg	agagtgaaa	ggggtaagag	aacagtaaa	aaaaaaatat	gacaaggagg	300
gccatgagga	tctacgattc	tagttacttt	cctcacgggt	gtcgcttgaa	gagcagggtg	360
agatcctcta	gaggttcaca	ggaatagcta	gcgttgcttc	ctggattttc	gggttccttt	420
ggcagtatac	agagtttgac	tcgagtgtga	tgtattcaag	actccactcc	agccacttta	480
accgcagttg	gggtagataa	aatgactggg	tagggtcctt	cccaggatgt	atctaaggat	540
ggggacttag	aaggaaaggga	cttgactaat	accatgtcac	cagggtgcaa	taattacttt	600
ccctcttctc	gggaacagggt	tcctgtaat	gttttaagaa	cttggtgata	tttgccaag	660
gaggtgatgt	ctgcaactaa	gctggccatc	tctcgggtcaa	gcacaagggtc	cttggttagg	720
aaggggccatc	catacagcat	tttgtatggg	ctaagtcctg	ctttttgggg	agagttttgg	780
attcttagta	aggctgtagg	caacagagca	ggccatgcaa	ggtgggtttc	ttgggttagc	840
ttttttaa	gtcgtttgag	tgcttcattc	attttcttga	cttttcctga	ggattgtggc	900
ctccacgcgc	agtgtaaagt	atattgtatg	cctaattgcct	gggatactcc	ctgggttact	960
gtagccttga	aaacggggcc	attgtcactc	tgtaagcctc	ggggaagtcc	gaatctggga	1020
attatttcat	gaattagtgc	ctttattaca	tcttggtcct	tttctgtcct	acaaagggaag	1080
gcctctgccc	aaccagtga	aatatctacc	cagactagta	gatactgaaa	tccctgagat	1140
ttgggcatgt	gggtaaaatc	tagttgccag	tcttctcctg	agtaatggcc	tgttctttgt	1200
tctcctgaag	gagctt					1216

<210> 8

<211> 976

<212> ADN

<213> Homo sapiens

<400> 8

agtataatg	gaataacttga	aagtaatccc	ctcactccag	gaactagtgc	tgagctggcc	60
aaactaatag	ccctcactcg	ggcactagaa	ttaggagaag	agaaaagggt	aaatatatat	120
acagactata	agtatgctta	cctagtccct	catgcccattg	cagcaatatg	gagagaaagg	180
gaattcctaa	cttccaaagg	aacacctatc	aaacatcagg	aagccattag	gatattatta	240
ttgggtgtac	agaaaccta	agaggtggca	gtcctacact	gctgggttca	tcagaaaaaa	300
aaggaaagg	aaatagaagg	gaactaccaa	gcagatattg	aagccaaaag	agccgcaagg	360
caggaccctc	cattagaaat	gcttatagaa	ggacccttag	tgtggggtaa	ccccctccag	420
gaaagcaatc	cccagtactc	agcaggagaa	ataaaatgga	gaacctcacg	aggacatact	480
ttcctcccct	caggatggct	agccaccaa	gaaggaaaaa	tgcttttgcc	tgacagtaac	540
caatggaaat	tacttaaaac	ccttcaccaa	acctttcact	taggattgat	agcaccatc	600

agatggccaa	attattatttt	actggatcag	gcctttttcaa	aactatcaag	caggtagtca	660
gggcctgtaa	agtgtgccaa	agaaataatc	tcctgcaactg	caagccatac	atttcaatcc	720
ctgtatcttt	aacctccttg	ttaagtttgt	ctcttcacaga	atcaaagctg	taaaactaca	780
aatgggttctt	caaattggagt	ctcagatgca	gtccatgact	aagatatacc	gcagccccct	840
ggagggggccc	tgctagccca	tgctccaatg	ttaatgacat	cgaaggcacc	cctcccgggg	900
aaatctcaac	tgcacaaccc	ctactatgtc	ccaattcagc	aggaagcagt	taaagcgggtc	960
atcggcccaac	ctcccc					976

<210> 9
 <211> 942
 <212> ADN
 <213> Homo sapiens

<400> 9						
agaggagaac	agcagcataa	gcggctggca	gaggtaggga	aagaccagca	agaagaaaag	60
agagaaaagag	aaagagaaaag	tcagagaaaag	agacagagag	aggaagagac	aaagagacag	120
aaagtcaaaag	aggtagtagt	cagaaacaga	gacaaaaaaa	aggagtcaga	aagagggaca	180
gacacagaaa	gtcaaaaaaa	aagttaagaa	gaaaggaaaa	gacaaagaag	aagtcgaaga	240
ggagaaaagag	agagatagaa	gtagttaaaga	aaaaaacagc	atatccccat	ccttttaaagc	300
cagggtaaat	ttctatctac	ccagccaagg	catattctac	ttatgtggat	cttcaaccca	360
tatctgcctc	tcagacagtt	tgcaagaaat	aatgaaatct	atccttactt	tacaatccca	420
aatagactct	ttggcagcag	tgactctcca	aaactgcaga	ggcctagacc	tcctcactgc	480
tgaaaaagga	ggacactaca	ccttcttagg	ggaagaatgt	tgtttttaca	ctaaccagtc	540
ggggatagta	tgagatgctg	cccggagttt	acaggaaaaag	gcttctgaaa	tcagacaacg	600
cctttcaaat	tcttatacca	acttctggag	ttaggcaaca	tggttctctc	cctttctagg	660
tcctgtggca	gccatcttgc	tgttactcgc	ctttggggccc	tgtattttta	accttcttgt	720
caaatttggt	tcctctagaa	tcgaggccat	caagctacag	atggctctac	aaatggaacc	780
ccaaaagagt	tcaactaaca	acttctaccg	aggacccctg	gatcaaccca	ctggcacttc	840
ccctggccta	gagagttccc	ctctgaagga	caccgcaact	tgagggccct	tctttgcccc	900
atccagcagg	agtagctaga	gtggtcatcg	gccaaattgc	ca		942

<210> 10
 <211> 1375
 <212> ADN
 <213> Homo sapiens

<400> 10						
ccccaatatt	ctctttctga	tggggaaaaa	tggccacctg	agggaagcac	aaattacaat	60
actatcctgc	agcttgatct	tttctgtaag	agggaaggca	aatggagtga	aataccttat	120
gtccaagctt	tcttttcatt	gagggagaat	acacaactat	gcaaagcttg	caattttacat	180
cccacaggag	gacccctcag	cttaccacca	tatcctagcc	tccctatagc	ttcccttctc	240
attgatgata	ctcctcctct	aatctcccc	gccagaaagg	aaataagcaa	agaaatctcc	300
aaaggtccac	aaaaaccccc	gggctatcgg	ttatgtcccc	ttcaagctgt	agggggaggg	360
gaatttggcc	caaccgggt	gcatgtcccc	ttctccctct	ctgattttaa	gcagatcagg	420
cagacctggg	gaagttttca	gatgatcctg	ataggtacat	agatgtccta	cagggtctag	480
ggcaaacctt	tgacctcact	tgagagagcg	tcagtctact	gttagatcaa	accctggcct	540
ttaatgaaaa	gaatgcggct	ttagctgcag	cctgagagtt	tgagataacc	tggtatccta	600
gtcaagtaaa	tgaagaatg	acagccgaag	aaagggacaa	cttccctact	ggtcagcaag	660
ccatccccag	tatggatccc	cactgggact	ttgactcaga	tcatggggac	tgagtcgta	720
aacatctgtt	gatctgtgtt	ctggaaggac	taaggagaat	tgggaaaaag	cccatgaatt	780
attcaatgat	atccaccata	acccagggaa	aggaagaaaa	tccttctgcc	ttcctcgagc	840
ggctacaaga	ggccttaaga	aaatatactc	ccctgtcacc	cgaatcactc	gagggtcaat	900
tgattctaaa	agataagttt	attaccctaat	cagccacaga	tatcaggaga	aagctccaaa	960
agcaagccct	gagccctgaa	caaaatctag	agacattatt	aaacctggca	accttggtgt	1020
tctataatag	ggaccaagag	gaacaggccc	aaaaggaaaa	gcgagatcag	agaaaggccg	1080
cagccttagt	catggccctc	agacaaacaa	accttggtgg	ttcagagagg	tcagaaaatg	1140

10

gagcaggcca	atcacctggt	acggcttggt	atcagtgccg	tttactagga	cacttttaaaa	1200
aagattgtcc	aataagaaac	aagctgcccc	ctcatccgtg	tccactatgc	cgaggcaatc	1260
actggaaggt	gcaactgccc	agaggatgaa	ggttccctgg	gttagaagcc	cccaaccaga	1320
tgatccaaca	acaggactga	gggtgcccgg	ggcaagcacc	agctcatgtc	atcac	1375

<210> 11
 <211> 944
 <212> ADN
 <213> Homo sapiens

<400> 11

acctaggagg	aactgtcttc	aggacaggac	tatagatgct	tcctcccagg	cgattaaggg	60
aaaaagacac	aatgggtatt	cagtaagtga	taaggaaact	cttgtagaag	cagagttagg	120
aaaattgcct	aataattggt	ctgctcaaat	gtgcgagctg	tttgcaactc	gccaaacctt	180
aaaagtatta	cagaatcagg	aagaagccat	ctataccaat	tctaagttaa	tatggactga	240
acgagaactt	attaatagca	aagaataatt	gaaatcccaa	acttacaagg	ttttcaacaa	300
aagcacagtt	tgctaaaagt	taactgtgta	acatgtatta	tcctactacc	acaaactctc	360
aaatgatttc	tcagacagtt	tgcaagaaac	aatgaaacct	atccttactc	tacaatccca	420
aatagactct	ttggcagcag	tgactctcca	aaaccaccaa	ggcctagacc	tcctcactgc	480
tgagaaagga	ggactctgca	ccttcttagg	ggaagattgt	tgtttttaca	ctaaccagtc	540
agggatagtg	tgagatgcca	cccagcgttt	acaggaaaag	gcttctgaaa	tcagacacaa	600
tgctttttcaa	accttatagc	aacctctgga	gttcggcgac	tggtttttcc	cctttctagg	660
tcctgtgaca	gccatcttgc	tattactcgc	cttcggggccc	tgtattttta	acctcctcgt	720
caaatttggt	tcctctagga	tcgaggccat	caagctacag	atggtcttac	aaatggaacc	780
ccaaatgagc	tcgactaaca	acttctactg	aggacccctg	gaccgaccca	ctggcccttt	840
aactggctta	aagagtttcc	ctctggagga	cactacaact	gcaggggccc	ttctttgccc	900
catccacagg	aagttagcta	gagcagtcac	cacccaattc	ccaa		944

<210> 12
 <211> 963
 <212> ADN
 <213> Homo sapiens

<400> 12

tacaggaacc	ccataatacg	tccttggtcaa	attctattca	gctccaactg	ctaggagtgg	60
cccatttgtc	ctgaaccctc	aaatcatggg	aatgagaaat	gaatttagac	tgaccacagc	120
ccttatgagt	tttcagctac	aggggtgtat	agaaccctga	taaggagtgt	tctttgtgtg	180
tggaagatcc	ttctatatatt	gcctccccac	caactggaca	ggaacttgta	ctttagccta	240
catagtacct	cctgtgactt	atccttttca	gaagaggcag	tagctgtgcc	cattcatgct	300
aagcttcagc	cgagagcaat	ctcactactt	cctctattgg	ctggtttagg	atttactacc	360
acctaggaag	tggaactcaca	gcctagatga	aatctctctc	caacttactc	aaatccagga	420
ccaaatagac	tcatttagcag	ctgtggttct	ccgaaccagt	gagcactaga	tctccaatct	480
cctcactgcc	gaaaggggag	gaacatgcct	ttttctgaac	aaggaaatgt	gtttttatgt	540
caataaatca	ggcatagtga	gagatggaat	taaatgactt	caggatagag	ctagcagact	600
acatgggtggg	acaaccgaaa	ctacctcagg	gttctcacag	cctgttctcc	actggettct	660
tccattttta	ggtcccttcc	ttatgattat	tctaggagta	acctttggcc	catgtctttt	720
cagttccttc	atcctttcgt	ttcttcctga	atagaatcaa	tgaaactaga	aatgttactg	780
cagatggaac	ctcagatgac	ttcaaccagc	acctattatc	aaggaccctt	aaaccagcct	840
gccggcccat	accggacgt	tgacacccaa	accacctctc	acgaggaaac	ctcagctaca	900
gaacccttcc	tatgccccta	ttcagcagga	agcaattaga	gtggtcatcc	tcccacaccc	960
caa						963

<210> 13
 <211> 1362
 <212> ADN
 <213> Homo sapiens

<400> 13
 ccacaatatc ctcttccagc aggagaacga tggccacctg agggaaagtat acactataat 60
 accatcctgc aactagatct gttttgtaaa caagaaggca agtggattta ggtaccatat 120
 gttcagacct ttttctcatt aagggatgat aaccacagat tgtgtaagac atgtaacctg 180
 caccacacag ggagtcctca aattctaccc ccataccag tcctccccac ggctcctcct 240
 actaatgccaa aaccctctct ggcttctaca gccaaaagg gaacaaataa aagagccttc 300
 agagagccaa gagacccac tggcccctgg ctatgtcctc ttcaggctgt agggagggaa 360
 tttggcccaa ccgagtaga tgttcccttt tctctctctg atctaaagca aattaaggca 420
 gacttgatg aaagtctctca gatgaccca atagatacgt agatggcctg ctgggtctgg 480
 gacaatcttt tgacctttcc tggagagaga tcatgttatt gcttgatcag acctaacctc 540
 taatgagaag aatgctgctt taacaggagc ccgagagtgt ggggatacct ggtacctcag 600
 ttaagtaagt gatagaatga catcagaaga gagcagtttc ctactggcca gcaagcagtc 660
 ccagtagtg atccccactg ggaccctgac tcggatcatg gggactggag tcacaaacat 720
 ttactgacct gtatcctaga agggttaagg agaactagga aaaagcccat gaactattca 780
 atgatgtcta ctataaccca agggaaaggaa gaaaacccta ttgccttcct caaaaggctg 840
 agggaggctt tgagaaaata tactcccctg tcaccagatt ccctcgaagg ccagttaatt 900
 ttaaaggaca aatttattac tcagtcagct gcagacatta ggaaaaagct ccaaaagtta 960
 gccttgggcc gagcaaaaatt tggaggcatc attaaacctg gcaacctcag tgttctatca 1020
 tagggaccaa gaggaacagg ccgaaaagga aaagcaggat aagagaaagg ctgcagattt 1080
 agtcatgccc tcagacaaac cttggcgggt caaagaggag aaaaaatgga gcaggccaat 1140
 caccacagcag ggcttattat cagtgcagtt tacaaggaca ctttaacaa gattgtccaa 1200
 agagaaataa gccgcccctc caccatgtc cactatgcca aggtgatcac tgggaaggcac 1260
 actgtccccg aggacaaagg ttctctgggc cagaagtccc caaccagatg atccagcaac 1320
 aggatggagg gtgcccgggg caagcaccag ctctgttgtt ca 1362

<210> 14
 <211> 945
 <212> ADN
 <213> Homo sapiens

<400> 14
 ttgcagatca atctcagact gctgtgctag caatgagtga ggcttcgtgg gcatgggacc 60
 ctctgagcca ggcatgggat ataatgtcct tgtgtgccat ttgctaagac tgttgggaata 120
 gcacagtatt agggtagggag tggcccagatt ttccagggtg tgtctgtcac cgcttccctt 180
 ggctaggaaa gagaattccc tgaccccttg ttcttcccag gtaaggcagt gcctcaccct 240
 gcttcagctc aactcaggt gactgcaccc actgtcctgc cccactgtc ggacaagccc 300
 cagtgcagtg aacctggtac ctgagttgga aatgcagaaa tcacctgtct tctgcgtcac 360
 tcacactggg agctgtagac tggagctgtt cctatttggc catcttggaa ccatctccca 420
 aatagactct ttggcagcag tgactctcca aaaccaccaa ggcctagacc tcctcattgc 480
 tgagaaagga ggactctgca ccttcttagg ggaggagtgt tgtttttata ctgaccagtc 540
 agggatggta cgagatgccaa cccgatgttt acaggaaaag gcttctgaaa tcacacaaca 600
 cctttcaaac tcttatacca acctctggag ttgggcaaca tggcttctcc cctttctcgg 660
 tccatttgca gccatcttgc tattactcgc cttcaggctg tgtattttta acctccttgt 720
 caaatttggt tcctctagaa ttgaggccgt caagctacag atggtcttac aaatgggacc 780
 ccaaatgag tcaactaaca acttctgcca aggacccctg gaccaacctg ctggcccttt 840
 cactggcctt aagagttccc ctctggagg cactacaact gcaggggccc ttctttgccc 900
 ctatccagca ggaagtagct agagcagtc taccccaatt cccaa 945

12

<210> 15
 <211> 939
 <212> ADN
 <213> Homo sapiens

<400> 15
 agagctacct tggcaagtac tctaggagta tgggaaaatg aaaacaacaa actcacacac 60
 cattttaaca tacacaatca ggtctgcccc cccagcaagg tatattcttt gtatgtggaa 120
 catcgacctt tatctgcctc cccactaact agacagccac ctgaatctta gtctttctaa 180
 gtcccaacag taacattgcc ccaggaaatc agaccatata agtatccctc aaagctcaag 240
 tctgtcagtg cagagccata caactaatac cctactttat agggtaagga atggctactg 300
 ctacaggaac cagaatagct agtttgttta cttcattatc ctactaccac acactctcaa 360
 atgatttctc agacagtttg caagaaataa cgaaatctat ccttactcta caatcccaaa 420
 tagactcctt ggcagcagtg accctccaaa acggctgagg cctagacctc ctactgcca 480
 agaaaggagg actctgcatt ttcttagggg aagagtgttt ttacactaac cagtcaggga 540
 cagtatgaga tgccactcgg agtttacagg aaaaggcttc tgaagtcaga caatgccttc 600
 caaactctat accaaactct ggagttgggc aacatggctt ctcctcttcc taggtcccgt 660
 gacagccatc ttgctattat ttgcctttga gccctgtatt tttaattctc ttttcaaat 720
 tgtttcctct ggatcgaggc catcgagcta cagatggtct tcacaaatgg aacccccaaa 780
 gagctcaact aacaacttct actgaggacc cctggactaa cctgctgacc ctttactgg 840
 cctgaagaat tcccctctgg aggacactac aactgcaggg ctccttcttt gcccttatcc 900
 agcaggaagt agctagagct gtcattgcct aattcctaa 939

<210> 16
 <211> 979
 <212> ADN
 <213> Homo sapiens

<400> 16
 agtgataatg gaatacttga aagtaatccc ctactcccc aggaactagt gctcagctgg 60
 cagaactaat agccctcact cgggtactag aatcaggaga aggaaaaagg gtaaatatat 120
 atacagactc taagtgtgct tacctagtcc tccatgcccc tgcagcaata tggagagaaa 180
 gggaattcct aacttccgag ggaacaccta tcaaacatca ggaagccatt aggaaattat 240
 tattggctgt acagaaacct aaagaggtgg cagttttaca ctgcccgggt catcagaaaag 300
 gaaaggaaaag ggaaatacaa gggagccacc aagttgatat tgaagtcaaa agagccacaa 360
 ggctggaccc tccattagaa atgcttatag gaggacccct agtatggggg aatccccctc 420
 gggaagccaa gcccagtagc tcagcaggag aaatagaata gggaacttca tgaggacata 480
 ctccctccc ctccagatgg ctagccacca ataaaggaaa aatacttttg cctgcagcta 540
 accaatagaa attacttaaa acccttcac aaaccttcca cttaggcatt gatagcacc 600
 atgagatggc caaattatta ttactggac caggcccttt caaaactatc aagcagatag 660
 tcagggcctg taaagtctgc caaagaaata atcccctgca ctgcaggcca tacatttcaa 720
 tccctgtatc tttaacctcc ttcttaaatt tgtctcttcc agaataaaag ctgtaaaatt 780
 acaaatagtt cttcaaatgg agccacagat gcagtccatg actaagatcc accacagacc 840
 cctggaccag cctgctagcc catgctccaa tgtaaatgac atcgaaggca ccccctcctg 900
 aggaaatctc aactgcacaa cccctactac gccccaattc agcagaaagc agttagagtg 960
 gtcacagacc aacctcccc 979

<210> 17
 <211> 1774
 <212> ADN
 <213> Homo sapiens

<400> 17
 catgctggta aaggaccgct agaatccagc agccaggacc actttctttg tggtaagaa 60
 aggtgggaaa acaggtgcag gactgctaca ctggtaagca taactaatcc gataagcaga 120
 ggtccatggg tggttacgca ccctggaaaag gaataagcat taggactata gaggacactc 180

taggactaat	gctcatcgga	aaatgactag	gggtactggc	atccctatgt	tcttttttca	240
gatgggaaat	gttcccccca	aggcagaaat	gcccctaaga	tgtattctgg	agaaatggga	300
ccaatctgac	catcagacac	taagaaagaa	atgacttata	ttcttctgca	gtaccacctg	360
gccacaatat	cttcttcaag	gggcagaaac	ctggcctcct	gaggggaagta	taaattataa	420
caccatctta	cagctagacc	tcttttgtag	aaaagaaggc	aaatggagtg	aagtgccata	480
tgtacaaaact	ttcttttcat	taagagataa	ctcccaatta	tgtaaaaagt	gtgattttatg	540
ccctacagga	agccctcaga	gtctacctcc	cgaccccagc	aagaccccaa	ctccttctcc	600
aactaataag	gacccccctt	caacccaaat	ggtccaaaag	gagatagaca	aaggggtaaa	660
caatgaacca	aagagtgccca	atattacacg	attatactcg	ctccaagcag	tgggaggaga	720
at ttggccca	gccagcgtgc	atgtaccttt	ttctctctca	gattttaaagc	aaattaaaat	780
agacctaggt	aaattctcag	ataaccctga	tggctatatt	gatgttttac	aaggggttagg	840
acaatccttt	gatctgacat	ggagagatat	aatgtttactg	ctaaatcaga	cactaaccctc	900
aaatgaaaaa	agtgtctgcc	taacagcagc	ctgagagttt	ggcgaactct	ggtatctcag	960
tcaggtcaat	gataggatga	caacagatga	aagagaatga	ttccccacag	gccagcaggc	1020
agttcccagt	gtagaccctc	attaggacac	agaatcagaa	cttggagatt	ggtgccacag	1080
acatttgcta	acttgcgctg	tagaaggact	aaggaaaact	aggaagaagc	ccatgaatta	1140
ttcaatgatg	tccccataaa	cacagggaaa	ggaagaaaat	cctactgcc	ttctggagag	1200
actaagggaa	ggattgagga	agcatacctc	cctgtcacct	gactctatta	aaggccaact	1260
aatctttaaag	gataagttta	tcaactcagtc	agctgcagag	attaagaaaa	aacttcaaaa	1320
gtatgcctta	ggcccagagc	aaaacttaga	aaccctactg	aacttggcaa	cctcagtttt	1380
ttataataga	gatcaggaag	agcaggggaa	tgggacaaat	gggataaaaa	aaaaaaaaaa	1440
aggtgactgc	tttagtcgtg	gccctcaggc	aaatggactt	tggaggctcc	agaaaaggga	1500
aaagctgagc	aaattgaatg	cctaacaggg	cttgcttcta	gtgtgggtcta	caaggacact	1560
ttaaaaaaga	ttgtccaagt	agaaacaagc	tgcccccttg	tccatgcccc	ttatgtcaag	1620
ggaatcactg	gaaggcccac	tgccccagga	gatgaaggtc	ctctgagtca	gaagccacta	1680
accagataat	ccagcagcag	gactgaggat	gcccagggca	agcgccagcc	catgccatca	1740
ccctcacaga	gccttgggta	tgcttgacca	ttga			1774

<210> 18
 <211> 938
 <212> ADN
 <213> Homo sapiens

<400> 18						
tgtaggaaga	actcccttca	ggacaggaca	atagatgggt	cctcccaggt	gattaaggaa	60
aaaagacaca	gtattcagta	agtgataagg	aaactcttgt	agaagcagag	ttagaaaaat	120
tgcctaataa	ttggtctgct	caaagtgtgt	agttgtttgc	actcagccaa	atcttaaagt	180
acttacagaa	tcaggaagca	gccatctata	ccaattctaa	gttaatatgg	actaaacgag	240
gtttttattag	tagcaaaagaa	aaattaaaaat	cccaaaactta	caaggttttc	aactaaagtt	300
tgccaaaagt	taacagtgtg	acatgtatta	tcctactatc	acacactctc	aaaggatttc	360
tcagacagtt	tgcaagaaat	aacgtaatct	atccttactc	tacagtccca	aatagactct	420
ttggtagcag	tgactctcca	aaactgccga	ggtctagacc	tcctcaatgc	tgagaaaggga	480
gaactctgca	ccttcttagg	ggaagagtgc	tgtttttaca	ctaaccagtc	agggatagta	540
tgagatactg	cctgacgttt	acaggaaaag	gcttctgaaa	tcagacaacg	cctttcaagc	600
tcttatacca	acctctggag	ttgggcaaca	tggcttctcc	ccttgctagg	tcctgtggca	660
gccatcttgc	tattacttgc	cttcggggccc	tgtattttta	acctccttgt	caaatttgtt	720
tcctctagga	tcaaggccat	caagctacag	atggtcttac	aaatggaacc	ccaaatgagc	780
tcaactaaca	acttctactg	aggacacctg	gactgaccca	ctggcccttt	cactggccta	840
aagagttccc	ttctggagga	cactacaact	gcaggggccc	gtcttcaccc	ctatccagca	900
ggaagtagct	agatcagtc	ttgcccatt	cccaacag			938

<210> 19
 <211> 1308
 <212> ADN
 <213> Homo sapiens

14

<400> 19

gatgcttgcc	ccaggcaccc	tcagtcctgt	tgttgatca	tctggtcggg	ggcttctggc	60
ccaaagaacc	tttgtcctct	gaggcagtgc	accttccagt	gattgcctca	gcattgtgga	120
catgggcaag	ggggcagctt	gtttctcact	ggacaatctt	ttttaagggtg	tccttccaaa	180
ccacactggt	aacaagccct	accaggtgat	tggcctgctc	tattttctgt	cctctctgaa	240
ccaccaaggt	ttgtctgtct	gagggtcatg	actaaggctg	tggcctttct	ctgatcttgc	300
ttttcctttt	tggcctgttc	ctcttggtac	ctattataga	acactgaggt	tgccaggttt	360
aacaatggct	ccagattttg	ttcagggcac	agggctcatt	ttggagcttt	ctcctgatat	420
ctgcagctga	ttgggtaata	aacttatctt	ttaggatcaa	ttgactctca	agagagttgg	480
gtgacagggg	agtatatattc	cttgaggcct	cccatagccg	ctctaggaag	gcagaaggat	540
tttcttcctt	tccctgagtt	ataaaagaca	tcattgaaca	actcatggac	tttttcccaa	600
ttctccgtag	tccttctaga	acacaggtca	gcagatgttt	acgactccag	tccccatgat	660
ctgagtctag	acaccagtgg	ggatccatac	tggggatggc	ctgctgactg	gtagggaatt	720
tgtccctttc	tttggctgtc	attctatcat	ttacttgact	aagataccaa	gtatctccaa	780
attctcaggc	tgcagctaaa	gctgcattct	tttcattaaa	ggccagggtt	tgatctaata	840
gcatgacatg	tctccaagtg	aggtcaaagg	tttggcctag	atccatagga	catcagagaa	900
ggagaagggg	acatacacct	gagttagcca	aattcccctc	cctctacagc	ttgaagggga	960
cataagcaat	agcctgggga	tttttgtggt	cctttggaga	tttctttgct	tgtttcttct	1020
tgggtggggg	agatttagagg	aggcttatca	gtaataggaa	ggggagctat	agggaggcta	1080
ggatatgggg	gtaagctgag	aggctatctt	gtgggatgta	aattgcaagc	tttgcatagt	1140
tgtggatttt	ccttacaatg	aaaataaagc	ttggacataa	ggtattttcac	tccattttgcc	1200
ttccctctta	cagaaaaggt	caagctgcag	gatagtactg	taattttatac	ttccttcagg	1260
tggccatttc	ttcccatcag	agagagaata	ctggggctgg	gccatagt		1308

<210> 20

<211> 711

<212> ADN

<213> Homo sapiens

<400> 20

actgagagac	aggactagct	ggatttccta	ggccgactaa	gaatccctaa	gcctagctgg	60
gaaggtgacc	acgtccacct	ttaaaccacg	ggcttgcaac	ttagctcaca	cctgaccaat	120
cagagagctc	actaaaatgc	taattaggca	aagacaggag	gtaaaagaaat	agccaatcat	180
ctattgcctg	agagcacagc	aggagggaca	acaatcgga	tataaaccca	ggcattcgag	240
ctggcaacag	cagccccctt	ttgggtccct	tccctttgta	tgggagctgt	tttcatgcta	300
tttcaactcta	ttaaactctt	caactgcact	cttctggtcc	atgtttctta	cggtctgagc	360
tgagcttttt	ctcaccttcc	accactgctg	tttgccacca	ccgcagacct	gccgctgact	420
cccatccctc	tggatcctgc	aggggtgtcc	ctgtgctcct	gatccagcga	ggcgccatt	480
gccgctccca	attgggctaa	aggcttgcca	ttgttcctgc	acggctaagt	gcctgggttt	540
gttctaattg	agctgaacac	tagtcaactg	gttccatggt	tctcttctgt	gacccacggc	600
ttctaataga	actataacac	ttaccacatg	gcccagatt	ccattccttg	gaatccgtga	660
ggccaagaac	tccaggtcag	agaatacgag	gcttgccacc	atcttggaag	c	711

<210> 21

<211> 711

<212> ADN

<213> Homo sapiens

<400> 21

actgagagac	aggactagct	ggatttccta	ggctgactaa	gaatccctaa	gcctagctgg	60
gaaggtgacc	acatccacct	ttaaaccacg	ggcttgcaac	ttagctcaca	cctgaccaat	120
cagagagctc	actaaaatgc	taattaggca	aagacaggag	gtaaaagaaat	agccaatcat	180
ctattgcctg	agagcacagc	aggagggaca	atgatcgga	tataaaccca	agtcttcgag	240
ccggcaacgg	caacccccct	tgggtccctt	ccctttgtat	gggagctctg	ttttcatgct	300
atttcaactg	attaaatctt	gcaactgcac	tcttctggct	catgtttctt	acggcttgag	360
ctgagctttc	gctcgccatc	caccactgct	gtttgccgcc	accgcagacc	cgccgctgac	420

15

```

tcccatccct ctggatcatg caggggtgtcc gctgtgctcc tgatccagcg aggcacccat 480
tgccgctccc aatcgggcta aaggcttgcc attgttcctg catgggctaag tgcctgggtt 540
catcctaatt gagctgaaca ctagtcaactg ggttccatgg ttctcttctg tgacccacag 600
cttctaataag agctataaca ctcaccgcat ggccaaggt tccattcctt gaatccataa 660
ggccaagaac cccagggtcag agaacacgag gcttgccacc atcttgggag c 711

```

<210> 22

<211> 2055

<212> ADN

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(2055)

<400> 22

```

ccc aag aca gcc aac tta gtt gca gac atc acc tcc tta gcc aaa tat 48
Pro Lys Thr Ala Asn Leu Val Ala Asp Ile Thr Ser Leu Ala Lys Tyr
  1                    5                    10                    15

caa caa gtt ctt aaa aca tta caa gga acc tat ccc tga gaa gag gga 96
Gln Gln Val Leu Lys Thr Leu Gln Gly Thr Tyr Pro Glu Glu Gly
                20                    25                    30

aaa gaa cta ttc cac cct tgt gac atg gta tta gtc aag tcc ctt ccc 144
Lys Glu Leu Phe His Pro Cys Asp Met Val Leu Val Lys Ser Leu Pro
                35                    40                    45

tct aat tcc cca tcc cta gat aca tcc tgg gaa gga ccc tac cca gtc 192
Ser Asn Ser Pro Ser Leu Asp Thr Ser Trp Glu Gly Pro Tyr Pro Val
                50                    55                    60

att tta tct acc cca act gcg gtt aaa gtg gct gga gtg gag tct tgg 240
Ile Leu Ser Thr Pro Thr Ala Val Lys Val Ala Gly Val Glu Ser Trp
                65                    70                    75                    80

ata cat cac act tga gtc aaa tcc tgg ata ctg cca aag gaa cct gaa 288
Ile His His Thr Val Lys Ser Trp Ile Leu Pro Lys Glu Pro Glu
                85                    90                    95

aat cca gga gac aac gct agc tat tcc tgt gaa cct cta gag gat ttg 336
Asn Pro Gly Asp Asn Ala Ser Tyr Ser Cys Glu Pro Leu Glu Asp Leu
                100                    105                    110

cgc ctg ctc ttc aaa caa caa cca gga gga aag taa cta aaa tca taa 384
Arg Leu Leu Phe Lys Gln Gln Pro Gly Gly Lys Leu Lys Ser
                115                    120                    125

atc ccc atg gcc ctc cct tat cat att ttt ctc ttt act gtt ctt tta 432
Ile Pro Met Ala Leu Pro Tyr His Ile Phe Leu Phe Thr Val Leu Leu
                130                    135                    140

ccc tct ttc act ctc act gca ccc cct cca tgc cgc tgt atg acc agt 480
Pro Ser Phe Thr Leu Thr Ala Pro Pro Pro Cys Arg Cys Met Thr Ser
145                    150                    155                    160

```

16

agc tcc cct tac caa gag ttt cta tgg aga atg cag cgt ccc gga aat	528
Ser Ser Pro Tyr Gln Glu Phe Leu Trp Arg Met Gln Arg Pro Gly Asn	
165 170 175	
att gat gcc cca tcg tat agg agt ctt tct aag gga acc ccc acc ttc	576
Ile Asp Ala Pro Ser Tyr Arg Ser Leu Ser Lys Gly Thr Pro Thr Phe	
180 185 190	
act gcc cac acc cat atg ccc cgc aac tgc tat cac tct gcc act ctt	624
Thr Ala His Thr His Met Pro Arg Asn Cys Tyr His Ser Ala Thr Leu	
195 200 205	
tgc atg cat gca aat act cat tat tgg aca gga aaa atg att aat cct	672
Cys Met His Ala Asn Thr His Tyr Trp Thr Gly Lys Met Ile Asn Pro	
210 215 220	
agt tgt cct gga gga ctt gga gtc act gtc tgt tgg act tac ttc acc	720
Ser Cys Pro Gly Gly Leu Gly Val Thr Val Cys Trp Thr Tyr Phe Thr	
225 230 235 240	
caa act ggt atg tct gat ggg ggt gga gtt caa gat cag gca aga gaa	768
Gln Thr Gly Met Ser Asp Gly Gly Gly Val Gln Asp Gln Ala Arg Glu	
245 250 255	
aaa cat gta aaa gaa gta atc tcc caa ctc acc cgg gta cat ggc acc	816
Lys His Val Lys Glu Val Ile Ser Gln Leu Thr Arg Val His Gly Thr	
260 265 270	
tct agc ccc tac aaa gga cta gat ctc tca aaa cta cat gaa acc ctc	864
Ser Ser Pro Tyr Lys Gly Leu Asp Leu Ser Lys Leu His Glu Thr Leu	
275 280 285	
cgt acc cat act cgc ctg gta agc cta ttt aat acc acc ctc act ggg	912
Arg Thr His Thr Arg Leu Val Ser Leu Phe Asn Thr Thr Leu Thr Gly	
290 295 300	
ctc cat gag gtc tcg gcc caa aac cct act aac tgt tgg ata tgc ctc	960
Leu His Glu Val Ser Ala Gln Asn Pro Thr Asn Cys Trp Ile Cys Leu	
305 310 315 320	
ccc ctg aac ttc agg cca tat gtt tca atc cct gta cct gaa caa tgg	1008
Pro Leu Asn Phe Arg Pro Tyr Val Ser Ile Pro Val Pro Glu Gln Trp	
325 330 335	
aac aac ttc agc aca gaa ata aac acc act tcc gtt tta gta gga cct	1056
Asn Asn Phe Ser Thr Glu Ile Asn Thr Thr Ser Val Leu Val Gly Pro	
340 345 350	
ctt gtt tcc aat ctg gaa ata acc cat acc tca aac ctc acc tgt gta	1104
Leu Val Ser Asn Leu Glu Ile Thr His Thr Ser Asn Leu Thr Cys Val	
355 360 365	
aaa ttt agc aat act aca tac aca acc aac tcc caa tgc atc agg tgg	1152
Lys Phe Ser Asn Thr Thr Tyr Thr Thr Asn Ser Gln Cys Ile Arg Trp	
370 375 380	

17

gta act cct ccc aca caa ata gtc tgc cta ccc tca gga ata ttt ttt	1200
Val Thr Pro Pro Thr Gln Ile Val Cys Leu Pro Ser Gly Ile Phe Phe	
385 390 395 400	
gtc tgt ggt acc tca gcc tat cgt tgt ttg aat ggc tct tca gaa tct	1248
Val Cys Gly Thr Ser Ala Tyr Arg Cys Leu Asn Gly Ser Ser Glu Ser	
405 410 415	
atg tgc ttc ctc tca ttc tta gtg ccc cct atg acc atc tac act gaa	1296
Met Cys Phe Leu Ser Phe Leu Val Pro Met Thr Ile Tyr Thr Glu	
420 425 430	
caa gat tta tac agt tat gtc ata tct aag ccc cgc aac aaa aga gta	1344
Gln Asp Leu Tyr Ser Tyr Val Ile Ser Lys Pro Arg Asn Lys Arg Val	
435 440 445	
ccc att ctt cct ttt gtt ata gga gca gga gtg cta ggt gca cta ggt	1392
Pro Ile Leu Pro Phe Val Ile Gly Ala Gly Val Leu Gly Ala Leu Gly	
450 455 460	
act ggc att ggc ggt atc aca acc tct act cag ttc tac tac aaa cta	1440
Thr Gly Ile Gly Gly Ile Thr Thr Ser Thr Gln Phe Tyr Tyr Lys Leu	
465 470 475 480	
tct caa gaa cta aat ggg gac atg gaa cgg gtc gcc gac tcc ctg gtc	1488
Ser Gln Glu Leu Asn Gly Asp Met Glu Arg Val Ala Asp Ser Leu Val	
485 490 495	
acc ttg caa gat caa ctt aac tcc cta gca gca gta gtc ctt caa aat	1536
Thr Leu Gln Asp Gln Leu Asn Ser Leu Ala Ala Val Val Leu Gln Asn	
500 505 510	
cga aga gct tta gac ttg cta acc gct gaa aga ggg gga acc tgt tta	1584
Arg Arg Ala Leu Asp Leu Leu Thr Ala Glu Arg Gly Gly Thr Cys Leu	
515 520 525	
ttt tta ggg gaa gaa tgc tgt tat tat gtt aat caa tcc gga atc gtc	1632
Phe Leu Gly Glu Glu Cys Cys Tyr Tyr Val Asn Gln Ser Gly Ile Val	
530 535 540	
act gag aaa gtt aaa gaa att cga gat cga ata caa cgt aga gca gag	1680
Thr Glu Lys Val Lys Glu Ile Arg Asp Arg Ile Gln Arg Arg Ala Glu	
545 550 555 560	
gag ctt cga aac act gga ccc tgg ggc ctc ctc agc caa tgg atg ccc	1728
Glu Leu Arg Asn Thr Gly Pro Trp Gly Leu Leu Ser Gln Trp Met Pro	
565 570 575	
tgg att ctc ccc ttc tta gga cct cta gca gct ata ata ttg cta ctc	1776
Trp Ile Leu Pro Phe Leu Gly Pro Leu Ala Ala Ile Ile Leu Leu Leu	
580 585 590	
ctc ttt gga ccc tgt atc ttt aac ctc ctt gtt aac ttt gtc tct tcc	1824
Leu Phe Gly Pro Cys Ile Phe Asn Leu Leu Val Asn Phe Val Ser Ser	
595 600 605	

18

aga atc gaa gct gta aaa cta caa atg gag ccc aag atg cag tcc aag 1872
 Arg Ile Glu Ala Val Lys Leu Gln Met Glu Pro Lys Met Gln Ser Lys
 610 615 620

act aag atc tac cgc aga ccc ctg gac cgg cct gct agc cca cga tct 1920
 Thr Lys Ile Tyr Arg Arg Pro Leu Asp Arg Pro Ala Ser Pro Arg Ser
 625 630 635 640

gat gtt aat gac atc aaa ggc acc cct cct gag gaa atc tca gct gca 1968
 Asp Val Asn Asp Ile Lys Gly Thr Pro Glu Glu Ile Ser Ala Ala
 645 650 655

caa cct cta cta cgc ccc aat tca gca gga agc agt tag agc ggt cgt 2016
 Gln Pro Leu Leu Arg Pro Asn Ser Ala Gly Ser Ser Ser Gly Arg
 660 665 670

cgg cca acc tcc cca aca gca ctt agg ttt tcc tgt tga 2055
 Arg Pro Thr Ser Pro Thr Ala Leu Arg Phe Ser Cys
 675 680 685

<210> 23
 <211> 28
 <212> PRT
 <213> Homo sapiens

<400> 23
 Pro Lys Thr Ala Asn Leu Val Ala Asp Ile Thr Ser Leu Ala Lys Tyr
 1 5 10 15
 Gln Gln Val Leu Lys Thr Leu Gln Gly Thr Tyr Pro
 20 25

<210> 24
 <211> 55
 <212> PRT
 <213> Homo sapiens

<400> 24
 Glu Glu Gly Lys Glu Leu Phe His Pro Cys Asp Met Val Leu Val Lys
 1 5 10 15
 Ser Leu Pro Ser Asn Ser Pro Ser Leu Asp Thr Ser Trp Glu Gly Pro
 20 25 30
 Tyr Pro Val Ile Leu Ser Thr Pro Thr Ala Val Lys Val Ala Gly Val
 35 40 45
 Glu Ser Trp Ile His His Thr
 50 55

<210> 25
 <211> 38
 <212> PRT
 <213> Homo sapiens

19

<400> 25

Val Lys Ser Trp Ile Leu Pro Lys Glu Pro Glu Asn Pro Gly Asp Asn
 1 5 10 15

Ala Ser Tyr Ser Cys Glu Pro Leu Glu Asp Leu Arg Leu Leu Phe Lys
 20 25 30

Gln Gln Pro Gly Gly Lys
 35

<210> 26

<211> 540

<212> PRT

<213> Homo sapiens

<400> 26

Ile Pro Met Ala Leu Pro Tyr His Ile Phe Leu Phe Thr Val Leu Leu
 1 5 10 15

Pro Ser Phe Thr Leu Thr Ala Pro Pro Pro Cys Arg Cys Met Thr Ser
 20 25 30

Ser Ser Pro Tyr Gln Glu Phe Leu Trp Arg Met Gln Arg Pro Gly Asn
 35 40 45

Ile Asp Ala Pro Ser Tyr Arg Ser Leu Ser Lys Gly Thr Pro Thr Phe
 50 55 60

Thr Ala His Thr His Met Pro Arg Asn Cys Tyr His Ser Ala Thr Leu
 65 70 75 80

Cys Met His Ala Asn Thr His Tyr Trp Thr Gly Lys Met Ile Asn Pro
 85 90 95

Ser Cys Pro Gly Gly Leu Gly Val Thr Val Cys Trp Thr Tyr Phe Thr
 100 105 110

Gln Thr Gly Met Ser Asp Gly Gly Gly Val Gln Asp Gln Ala Arg Glu
 115 120 125

Lys His Val Lys Glu Val Ile Ser Gln Leu Thr Arg Val His Gly Thr
 130 135 140

Ser Ser Pro Tyr Lys Gly Leu Asp Leu Ser Lys Leu His Glu Thr Leu
 145 150 155 160

Arg Thr His Thr Arg Leu Val Ser Leu Phe Asn Thr Thr Leu Thr Gly
 165 170 175

Leu His Glu Val Ser Ala Gln Asn Pro Thr Asn Cys Trp Ile Cys Leu
 180 185 190

Pro Leu Asn Phe Arg Pro Tyr Val Ser Ile Pro Val Pro Glu Gln Trp
 195 200 205

Asn Asn Phe Ser Thr Glu Ile Asn Thr Thr Ser Val Leu Val Gly Pro
 210 215 220

Leu Val Ser Asn Leu Glu Ile Thr His Thr Ser Asn Leu Thr Cys Val
 225 230 235 240
 Lys Phe Ser Asn Thr Thr Tyr Thr Thr Asn Ser Gln Cys Ile Arg Trp
 245 250 255
 Val Thr Pro Pro Thr Gln Ile Val Cys Leu Pro Ser Gly Ile Phe Phe
 260 265 270
 Val Cys Gly Thr Ser Ala Tyr Arg Cys Leu Asn Gly Ser Ser Glu Ser
 275 280 285
 Met Cys Phe Leu Ser Phe Leu Val Pro Pro Met Thr Ile Tyr Thr Glu
 290 295 300
 Gln Asp Leu Tyr Ser Tyr Val Ile Ser Lys Pro Arg Asn Lys Arg Val
 305 310 315 320
 Pro Ile Leu Pro Phe Val Ile Gly Ala Gly Val Leu Gly Ala Leu Gly
 325 330 335
 Thr Gly Ile Gly Gly Ile Thr Thr Ser Thr Gln Phe Tyr Tyr Lys Leu
 340 345 350
 Ser Gln Glu Leu Asn Gly Asp Met Glu Arg Val Ala Asp Ser Leu Val
 355 360 365
 Thr Leu Gln Asp Gln Leu Asn Ser Leu Ala Ala Val Val Leu Gln Asn
 370 375 380
 Arg Arg Ala Leu Asp Leu Leu Thr Ala Glu Arg Gly Gly Thr Cys Leu
 385 390 395 400
 Phe Leu Gly Glu Glu Cys Cys Tyr Tyr Val Asn Gln Ser Gly Ile Val
 405 410 415
 Thr Glu Lys Val Lys Glu Ile Arg Asp Arg Ile Gln Arg Arg Ala Glu
 420 425 430
 Glu Leu Arg Asn Thr Gly Pro Trp Gly Leu Leu Ser Gln Trp Met Pro
 435 440 445
 Trp Ile Leu Pro Phe Leu Gly Pro Leu Ala Ala Ile Ile Leu Leu Leu
 450 455 460
 Leu Phe Gly Pro Cys Ile Phe Asn Leu Leu Val Asn Phe Val Ser Ser
 465 470 475 480
 Arg Ile Glu Ala Val Lys Leu Gln Met Glu Pro Lys Met Gln Ser Lys
 485 490 495
 Thr Lys Ile Tyr Arg Arg Pro Leu Asp Arg Pro Ala Ser Pro Arg Ser
 500 505 510
 Asp Val Asn Asp Ile Lys Gly Thr Pro Pro Glu Glu Ile Ser Ala Ala
 515 520 525

21

Gln Pro Leu Leu Arg Pro Asn Ser Ala Gly Ser Ser
 530 535 540

<210> 27
 <211> 15
 <212> PRT
 <213> Homo sapiens

<400> 27
 Ser Gly Arg Arg Pro Thr Ser Pro Thr Ala Leu Arg Phe Ser Cys
 1 5 10 15

<210> 28
 <211> 1080
 <212> ADN
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (1)..(1080)

<400> 28
 acc tct ttt gta gaa aag gca aat gga gtg aag tgc cat aag tac aaa 48
 Thr Ser Phe Val Glu Lys Ala Asn Gly Val Lys Cys His Lys Tyr Lys
 1 5 10 15

ctt tct ttt cat taa gag aca act cac aat tat gta aaa agt gtg att 96
 Leu Ser Phe His Glu Thr Thr His Asn Tyr Val Lys Ser Val Ile
 20 25 30

tat gcc cta cag gaa gcc ttc aga gtc tac ctc cct atc cca gca tcc 144
 Tyr Ala Leu Gln Glu Ala Phe Arg Val Tyr Leu Pro Ile Pro Ala Ser
 35 40 45

ccg act cct tcc cca act aat aag gac ccc cct tca acc caa atg gtc 192
 Pro Thr Pro Ser Pro Thr Asn Lys Asp Pro Pro Ser Thr Gln Met Val
 50 55 60

caa aag gag ata gac aaa agg gta aac agt gaa cca aag agt gcc aat 240
 Gln Lys Glu Ile Asp Lys Arg Val Asn Ser Glu Pro Lys Ser Ala Asn
 65 70 75 80

att ccc caa tta tga ccc ctc caa gca gtg gga gga aga gaa ttc ggc 288
 Ile Pro Gln Leu Pro Leu Gln Ala Val Gly Gly Arg Glu Phe Gly
 85 90 95

cca gcc aga gtg cat gtg cct ttt tct ctc cca gac tta aag caa ata 336
 Pro Ala Arg Val His Val Pro Phe Ser Leu Pro Asp Leu Lys Gln Ile
 100 105 110

aaa aca gac tta ggt aaa ttc tca gat aac cct gat ggc tat att gat 384
 Lys Thr Asp Leu Gly Lys Phe Ser Asp Asn Pro Asp Gly Tyr Ile Asp
 115 120 125

22

gtt tta caa ggg tta gga caa ttc ttt gat ctg aca tgg aga gat ata	432
Val Leu Gln Gly Leu Gly Gln Phe Phe Asp Leu Thr Trp Arg Asp Ile	
130 135 140	
atg tca ctg cta aat cag aca cta acc cca aat gag aga agt gcc acc	480
Met Ser Leu Leu Asn Gln Thr Leu Thr Pro Asn Glu Arg Ser Ala Thr	
145 150 155 160	
ata act gca gcc tga gag ttt ggc gat ctc tgg tat ctc agt cag gtc	528
Ile Thr Ala Ala 165 Glu Phe Gly Asp Leu Trp Tyr Leu Ser Gln Val	
170 175	
aat gat agg atg aca aca gag gaa aga gaa tga ttc ccc aca ggc cag	576
Asn Asp Arg Met Thr Thr Glu Glu Arg Glu Phe Pro Thr Gly Gln	
180 185 190	
cag gca gtt ccc agt cta gac cct cat tgg gac aca gaa tca gaa cat	624
Gln Ala Val Pro Ser Leu Asp Pro His Trp Asp Thr Glu Ser Glu His	
195 200 205	
gga gat tgg tgc tgc aga cat ttg cta act tgt gtg cta gaa gga cta	672
Gly Asp Trp Cys Cys Arg His Leu Leu Thr Cys Val Leu Glu Gly Leu	
210 215 220	
agg aaa act agg aag aag tct atg aat tac tca atg atg tcc acc ata	720
Arg Lys Thr Arg Lys Lys Ser Met Asn Tyr Ser Met Met Ser Thr Ile	
225 230 235 240	
aca cag gga agg gaa gaa aat cct act gcc ttt ctg gag aga cta agg	768
Thr Gln Gly Arg Glu Glu Asn Pro Thr Ala Phe Leu Glu Arg Leu Arg	
245 250 255	
gag gca ttg agg aag cgt gcc tct ctg tca cct gac tct tct gaa ggc	816
Glu Ala Leu Arg Lys Arg Ala Ser Leu Ser Pro Asp Ser Ser Glu Gly	
260 265 270	
caa cta atc tta aag cgt aag ttt atc act cag tca gct gca gac att	864
Gln Leu Ile Leu Lys Arg Lys Phe Ile Thr Gln Ser Ala Ala Asp Ile	
275 280 285	
aga aaa aaa ctt caa aag tct gcc gta ggc ccg gag caa aac tta gaa	912
Arg Lys Lys Leu Gln Lys Ser Ala Val Gly Pro Glu Gln Asn Leu Glu	
290 295 300	
acc cta ttg aac ttg gca acc tcg gtt ttt tat aat aga gat cag gag	960
Thr Leu Leu Asn Leu Ala Thr Ser Val Phe Tyr Asn Arg Asp Gln Glu	
305 310 315 320	
gag cag gcg gaa cag gac aaa cgg gat taa aaa aaa ggc cac cgc ttt	1008
Glu Gln Ala Glu Gln Asp Lys Arg Asp 330 Lys Lys Gly His Arg Phe	
325 335	
agt cat gac cct cag gca agt gga ctt tgg agg ctc tgg aaa agg gaa	1056
Ser His Asp Pro Gln Ala Ser Gly Leu Trp Arg Leu Trp Lys Arg Glu	
340 345 350	

23

aag ctg ggc aaa ttg aat gcc taa
 Lys Leu Gly Lys Leu Asn Ala
 355 360

1080

<210> 29
 <211> 20
 <212> PRT
 <213> Homo sapiens

<400> 29
 Thr Ser Phe Val Glu Lys Ala Asn Gly Val Lys Cys His Lys Tyr Lys
 1 5 10 15
 Leu Ser Phe His
 20

<210> 30
 <211> 63
 <212> PRT
 <213> Homo sapiens

<400> 30
 Glu Thr Thr His Asn Tyr Val Lys Ser Val Ile Tyr Ala Leu Gln Glu
 1 5 10 15
 Ala Phe Arg Val Tyr Leu Pro Ile Pro Ala Ser Pro Thr Pro Ser Pro
 20 25 30
 Thr Asn Lys Asp Pro Pro Ser Thr Gln Met Val Gln Lys Glu Ile Asp
 35 40 45
 Lys Arg Val Asn Ser Glu Pro Lys Ser Ala Asn Ile Pro Gln Leu
 50 55 60

<210> 31
 <211> 79
 <212> PRT
 <213> Homo sapiens

<400> 31
 Pro Leu Gln Ala Val Gly Gly Arg Glu Phe Gly Pro Ala Arg Val His
 1 5 10 15
 Val Pro Phe Ser Leu Pro Asp Leu Lys Gln Ile Lys Thr Asp Leu Gly
 20 25 30
 Lys Phe Ser Asp Asn Pro Asp Gly Tyr Ile Asp Val Leu Gln Gly Leu
 35 40 45
 Gly Gln Phe Phe Asp Leu Thr Trp Arg Asp Ile Met Ser Leu Leu Asn
 50 55 60
 Gln Thr Leu Thr Pro Asn Glu Arg Ser Ala Thr Ile Thr Ala Ala
 65 70 75

24

<210> 32
 <211> 21
 <212> PRT
 <213> Homo sapiens

<400> 32
 Glu Phe Gly Asp Leu Trp Tyr Leu Ser Gln Val Asn Asp Arg Met Thr
 1 5 10 15
 Thr Glu Glu Arg Glu
 20

<210> 33
 <211> 142
 <212> PRT
 <213> Homo sapiens

<400> 33
 Phe Pro Thr Gly Gln Gln Ala Val Pro Ser Leu Asp Pro His Trp Asp
 1 5 10 15
 Thr Glu Ser Glu His Gly Asp Trp Cys Cys Arg His Leu Leu Thr Cys
 20 25 30
 Val Leu Glu Gly Leu Arg Lys Thr Arg Lys Lys Ser Met Asn Tyr Ser
 35 40 45
 Met Met Ser Thr Ile Thr Gln Gly Arg Glu Glu Asn Pro Thr Ala Phe
 50 55 60
 Leu Glu Arg Leu Arg Glu Ala Leu Arg Lys Arg Ala Ser Leu Ser Pro
 65 70 75 80
 Asp Ser Ser Glu Gly Gln Leu Ile Leu Lys Arg Lys Phe Ile Thr Gln
 85 90 95
 Ser Ala Ala Asp Ile Arg Lys Lys Leu Gln Lys Ser Ala Val Gly Pro
 100 105 110
 Glu Gln Asn Leu Glu Thr Leu Leu Asn Leu Ala Thr Ser Val Phe Tyr
 115 120 125
 Asn Arg Asp Gln Glu Glu Gln Ala Glu Gln Asp Lys Arg Asp
 130 135 140

<210> 34
 <211> 29
 <212> PRT
 <213> Homo sapiens

<400> 34
 Lys Lys Gly His Arg Phe Ser His Asp Pro Gln Ala Ser Gly Leu Trp
 1 5 10 15

25

Arg Leu Trp Lys Arg Glu Lys Leu Gly Lys Leu Asn Ala
 20 25

<210> 35

<211> 685

<212> PRT

<213> Homo sapiens

<400> 35

Pro Lys Thr Ala Asn Leu Val Ala Asp Ile Thr Ser Leu Ala Lys Tyr
 1 5 10 15

Gln Gln Val Leu Lys Thr Leu Gln Gly Thr Tyr Pro Xaa Glu Glu Gly
 20 25 30

Lys Glu Leu Phe His Pro Cys Asp Met Val Leu Val Lys Ser Leu Pro
 35 40 45

Ser Asn Ser Pro Ser Leu Asp Thr Ser Trp Glu Gly Pro Tyr Pro Val
 50 55 60

Ile Leu Ser Thr Pro Thr Ala Val Lys Val Ala Gly Val Glu Ser Trp
 65 70 75 80

Ile His His Thr Xaa Val Lys Ser Trp Ile Leu Pro Lys Glu Pro Glu
 85 90 95

Asn Pro Gly Asp Asn Ala Ser Tyr Ser Cys Glu Pro Leu Glu Asp Leu
 100 105 110

Arg Leu Leu Phe Lys Gln Gln Pro Gly Gly Lys Xaa Leu Lys Ser Xaa
 115 120 125

Ile Pro Met Ala Leu Pro Tyr His Ile Phe Leu Phe Thr Val Leu Leu
 130 135 140

Pro Ser Phe Thr Leu Thr Ala Pro Pro Pro Cys Arg Cys Met Thr Ser
 145 150 155 160

Ser Ser Pro Tyr Gln Glu Phe Leu Trp Arg Met Gln Arg Pro Gly Asn
 165 170 175

Ile Asp Ala Pro Ser Tyr Arg Ser Leu Ser Lys Gly Thr Pro Thr Phe
 180 185 190

Thr Ala His Thr His Met Pro Arg Asn Cys Tyr His Ser Ala Thr Leu
 195 200 205

Cys Met His Ala Asn Thr His Tyr Trp Thr Gly Lys Met Ile Asn Pro
 210 215 220

Ser Cys Pro Gly Gly Leu Gly Val Thr Val Cys Trp Thr Tyr Phe Thr
 225 230 235 240

Gln Thr Gly Met Ser Asp Gly Gly Gly Val Gln Asp Gln Ala Arg Glu
 245 250 255

Lys His Val Lys Glu Val Ile Ser Gln Leu Thr Arg Val His Gly Thr
 260 265 270
 Ser Ser Pro Tyr Lys Gly Leu Asp Leu Ser Lys Leu His Glu Thr Leu
 275 280 285
 Arg Thr His Thr Arg Leu Val Ser Leu Phe Asn Thr Thr Leu Thr Gly
 290 295 300
 Leu His Glu Val Ser Ala Gln Asn Pro Thr Asn Cys Trp Ile Cys Leu
 305 310 315 320
 Pro Leu Asn Phe Arg Pro Tyr Val Ser Ile Pro Val Pro Glu Gln Trp
 325 330 335
 Asn Asn Phe Ser Thr Glu Ile Asn Thr Thr Ser Val Leu Val Gly Pro
 340 345 350
 Leu Val Ser Asn Leu Glu Ile Thr His Thr Ser Asn Leu Thr Cys Val
 355 360 365
 Lys Phe Ser Asn Thr Thr Tyr Thr Thr Asn Ser Gln Cys Ile Arg Trp
 370 375 380
 Val Thr Pro Pro Thr Gln Ile Val Cys Leu Pro Ser Gly Ile Phe Phe
 385 390 395 400
 Val Cys Gly Thr Ser Ala Tyr Arg Cys Leu Asn Gly Ser Ser Glu Ser
 405 410 415
 Met Cys Phe Leu Ser Phe Leu Val Pro Pro Met Thr Ile Tyr Thr Glu
 420 425 430
 Gln Asp Leu Tyr Ser Tyr Val Ile Ser Lys Pro Arg Asn Lys Arg Val
 435 440 445
 Pro Ile Leu Pro Phe Val Ile Gly Ala Gly Val Leu Gly Ala Leu Gly
 450 455 460
 Thr Gly Ile Gly Gly Ile Thr Thr Ser Thr Gln Phe Tyr Tyr Lys Leu
 465 470 475 480
 Ser Gln Glu Leu Asn Gly Asp Met Glu Arg Val Ala Asp Ser Leu Val
 485 490 495
 Thr Leu Gln Asp Gln Leu Asn Ser Leu Ala Ala Val Val Leu Gln Asn
 500 505 510
 Arg Arg Ala Leu Asp Leu Leu Thr Ala Glu Arg Gly Gly Thr Cys Leu
 515 520 525
 Phe Leu Gly Glu Glu Cys Cys Tyr Tyr Val Asn Gln Ser Gly Ile Val
 530 535 540
 Thr Glu Lys Val Lys Glu Ile Arg Asp Arg Ile Gln Arg Arg Ala Glu
 545 550 555 560

27

Glu Leu Arg Asn Thr Gly Pro Trp Gly Leu Leu Ser Gln Trp Met Pro
 565 570 575
 Trp Ile Leu Pro Phe Leu Gly Pro Leu Ala Ala Ile Ile Leu Leu Leu
 580 585 590
 Leu Phe Gly Pro Cys Ile Phe Asn Leu Leu Val Asn Phe Val Ser Ser
 595 600 605
 Arg Ile Glu Ala Val Lys Leu Gln Met Glu Pro Lys Met Gln Ser Lys
 610 615 620
 Thr Lys Ile Tyr Arg Arg Pro Leu Asp Arg Pro Ala Ser Pro Arg Ser
 625 630 635 640
 Asp Val Asn Asp Ile Lys Gly Thr Pro Pro Glu Glu Ile Ser Ala Ala
 645 650 655
 Gln Pro Leu Leu Arg Pro Asn Ser Ala Gly Ser Ser Xaa Ser Gly Arg
 660 665 670
 Arg Pro Thr Ser Pro Thr Ala Leu Arg Phe Ser Cys Xaa
 675 680 685

<210> 36
 <211> 360
 <212> PRT
 <213> Homo sapiens

<400> 36
 Thr Ser Phe Val Glu Lys Ala Asn Gly Val Lys Cys His Lys Tyr Lys
 1 5 10 15
 Leu Ser Phe His Xaa Glu Thr Thr His Asn Tyr Val Lys Ser Val Ile
 20 25 30
 Tyr Ala Leu Gln Glu Ala Phe Arg Val Tyr Leu Pro Ile Pro Ala Ser
 35 40 45
 Pro Thr Pro Ser Pro Thr Asn Lys Asp Pro Pro Ser Thr Gln Met Val
 50 55 60
 Gln Lys Glu Ile Asp Lys Arg Val Asn Ser Glu Pro Lys Ser Ala Asn
 65 70 75 80
 Ile Pro Gln Leu Xaa Pro Leu Gln Ala Val Gly Gly Arg Glu Phe Gly
 85 90 95
 Pro Ala Arg Val His Val Pro Phe Ser Leu Pro Asp Leu Lys Gln Ile
 100 105 110
 Lys Thr Asp Leu Gly Lys Phe Ser Asp Asn Pro Asp Gly Tyr Ile Asp
 115 120 125
 Val Leu Gln Gly Leu Gly Gln Phe Phe Asp Leu Thr Trp Arg Asp Ile
 130 135 140

28

Met Ser Leu Leu Asn Gln Thr Leu Thr Pro Asn Glu Arg Ser Ala Thr
 145 150 155 160

Ile Thr Ala Ala Xaa Glu Phe Gly Asp Leu Trp Tyr Leu Ser Gln Val
 165 170 175

Asn Asp Arg Met Thr Thr Glu Glu Arg Glu Xaa Phe Pro Thr Gly Gln
 180 185 190

Gln Ala Val Pro Ser Leu Asp Pro His Trp Asp Thr Glu Ser Glu His
 195 200 205

Gly Asp Trp Cys Cys Arg His Leu Leu Thr Cys Val Leu Glu Gly Leu
 210 215 220

Arg Lys Thr Arg Lys Lys Ser Met Asn Tyr Ser Met Met Ser Thr Ile
 225 230 235 240

Thr Gln Gly Arg Glu Glu Asn Pro Thr Ala Phe Leu Glu Arg Leu Arg
 245 250 255

Glu Ala Leu Arg Lys Arg Ala Ser Leu Ser Pro Asp Ser Ser Glu Gly
 260 265 270

Gln Leu Ile Leu Lys Arg Lys Phe Ile Thr Gln Ser Ala Ala Asp Ile
 275 280 285

Arg Lys Lys Leu Gln Lys Ser Ala Val Gly Pro Glu Gln Asn Leu Glu
 290 295 300

Thr Leu Leu Asn Leu Ala Thr Ser Val Phe Tyr Asn Arg Asp Gln Glu
 305 310 315 320

Glu Gln Ala Glu Gln Asp Lys Arg Asp Xaa Lys Lys Gly His Arg Phe
 325 330 335

Ser His Asp Pro Gln Ala Ser Gly Leu Trp Arg Leu Trp Lys Arg Glu
 340 345 350

Lys Leu Gly Lys Leu Asn Ala Xaa
 355 360

<210> 37
 <211> 26
 <212> ADN
 <213> Homo sapiens

<400> 37
 ggacataga ggacactcca ggacta

26

<210> 38
 <211> 25
 <212> ADN
 <213> Homo sapiens

<400> 38 cctcagtcct gctgctggat catct	25
<210> 39 <211> 27 <212> ADN <213> Homo sapiens	
<400> 39 cctccaagca gtgggaggaa gagaatt	27
<210> 40 <211> 28 <212> ADN <213> Homo sapiens	
<400> 40 ccttcctgt gttattgtgg acatcatt	28
<210> 41 <211> 30 <212> ADN <213> Homo sapiens	
<400> 41 ggaagaagtc tatgaattat tcaatgatgt	30
<210> 42 <211> 27 <212> ADN <213> Homo sapiens	
<400> 42 gggacacaga atcagaacat ggagatt	27
<210> 43 <211> 27 <212> ADN <213> Homo sapiens	
<400> 43 gccttcagaa gagtcaggtg acagaga	27
<210> 44 <211> 25 <212> ADN <213> Homo sapiens	
<400> 44 gagcctccaa agtccacttg cctga	25

<210> 45
<211> 29
<212> ADN
<213> Homo sapiens

<400> 45
gatttcagta tctactagtc tgggtagat 29

<210> 46
<211> 27
<212> ADN
<213> Homo sapiens

<400> 46
ctaggaaatc cagctagtcc tgtctca 27

<210> 47
<211> 28
<212> ADN
<213> Homo sapiens

<400> 47
ccaagacagc caacttagtt gcagacat 28

<210> 48
<211> 28
<212> ADN
<213> Homo sapiens

<400> 48
ggacgctgca ttctccatag aaactctt 28

<210> 49
<211> 29
<212> ADN
<213> Homo sapiens

<400> 49
gcaatactac atacacaacc aactcccaa 29

<210> 50
<211> 26
<212> ADN
<213> Homo sapiens

<400> 50
gggggaggca tatccaacag ttagta 26

31

<210> 51
<211> 30
<212> ADN
<213> Homo sapiens

<400> 51
ccatctacac tgaacaagat ttatacactt 30

<210> 52
<211> 28
<212> ADN
<213> Homo sapiens

<400> 52
aatgccagta cctagtgcac ctagcact 28

<210> 53
<211> 31
<212> ADN
<213> Homo sapiens

<400> 53
cgaatacaac gtagagcaga ggagcttcga a 31

<210> 54
<211> 28
<212> ADN
<213> Homo sapiens

<400> 54
agcccaagat gcagtccaag actaagat 28

<210> 55
<211> 27
<212> ADN
<213> Homo sapiens

<400> 55
gcgtagtaga ggttgtgcag ctgagat 27

<210> 56
<211> 27
<212> ADN
<213> Homo sapiens

<400> 56
cccttaccaa gagtttctat ggagaat 27

<210> 57
<211> 27
<212> ADN
<213> Homo sapiens

32

<400> 57
accgctctaa ctgcttcctg ctgaatt

27

<210> 58
<211> 420
<212> PRT
<213> Homo sapiens

<400> 58
Thr Ser Phe Val Glu Lys Ala Asn Gly Val Lys Cys His Lys Tyr Lys
1 5 10 15
Leu Ser Phe His Xaa Glu Thr Thr His Asn Tyr Val Lys Ser Val Ile
20 25 30
Tyr Ala Leu Gln Glu Ala Phe Arg Val Tyr Leu Pro Ile Leu Pro Ala
35 40 45
Ser Pro Thr Pro Ser Pro Thr Asn Lys Asp Pro Pro Ser Thr Gln Met
50 55 60
Val Gln Lys Glu Ile Asp Lys Arg Val Asn Ser Glu Pro Lys Ser Ala
65 70 75 80
Asn Ile Pro Gln Leu Xaa Pro Leu Gln Ala Val Gly Gly Arg Glu Phe
85 90 95
Gly Pro Ala Arg Val His Val Pro Phe Ser Leu Pro Asp Leu Lys Gln
100 105 110
Ile Lys Thr Asp Leu Gly Lys Phe Ser Asp Asn Pro Asp Gly Tyr Ile
115 120 125
Asp Val Leu Gln Gly Leu Gly Gln Phe Phe Asp Leu Thr Trp Arg Asp
130 135 140
Ile Met Ser Leu Leu Asn Gln Thr Leu Thr Pro Asn Glu Arg Ser Ala
145 150 155 160
Thr Ile Thr Ala Ala Xaa Glu Phe Gly Asp Leu Trp Tyr Leu Ser Gln
165 170 175
Val Asn Asp Arg Met Thr Thr Glu Glu Arg Glu Xaa Phe Pro Thr Gly
180 185 190
Gln Gln Ala Val Pro Ser Leu Asp Pro His Trp Asp Thr Glu Ser Glu
195 200 205
His Gly Asp Trp Cys Cys Arg His Leu Leu Thr Cys Val Leu Glu Gly
210 215 220
Leu Arg Lys Thr Arg Lys Lys Ser Met Asn Tyr Ser Met Met Ser Thr
225 230 235 240
Ile Thr Gln Gly Arg Glu Glu Asn Pro Thr Ala Phe Leu Glu Arg Leu
245 250 255

33

Arg Glu Ala Leu Arg Lys Arg Ala Ser Leu Ser Pro Asp Ser Ser Glu
 260 265 270
 Gly Gln Leu Ile Leu Lys Arg Lys Phe Ile Thr Gln Ser Ala Ala Asp
 275 280 285
 Ile Arg Lys Lys Leu Gln Lys Ser Ala Val Gly Pro Glu Gln Asn Leu
 290 295 300
 Glu Thr Leu Leu Asn Leu Ala Thr Ser Val Phe Tyr Asn Arg Asp Gln
 305 310 315 320
 Glu Glu Gln Ala Glu Gln Asp Lys Arg Asp Xaa Lys Lys Gly His Arg
 325 330 335
 Phe Ser His Asp Pro Gln Ala Ser Gly Leu Trp Arg Leu Trp Lys Arg
 340 345 350
 Glu Lys Leu Gly Lys Leu Asn Ala Xaa Xaa Gly Leu Leu Pro Val Arg
 355 360 365
 Ser Thr Arg Thr Leu Xaa Lys Arg Leu Ser Lys Xaa Lys Xaa Ala Ala
 370 375 380
 Pro Ser Ser Met Pro Leu Ile Ser Arg Glu Ser Leu Glu Gly Pro Leu
 385 390 395 400
 Pro Gln Gly Thr Lys Val Leu Xaa Val Arg Ser His Xaa Pro Asp Ser
 405 410 415
 Ser Ser Arg Thr
 420

<210> 59
 <211> 32
 <212> ADN
 <213> Homo sapiens

<400> 59
 taaactacaa atggttcttc aaatggagcc ca

32

<210> 60
 <211> 32
 <212> ADN
 <213> Homo sapiens

<400> 60
 gatgcagtcc aagatgcagt ccatgactaa ga

32

<210> 61
 <211> 1740
 <212> ADN
 <213> Homo sapiens

<400> 61

```

aggttggtc acaaccgctc ttaactgctt catgctgaat tggggcatag taggggtcgt 60
gcagttgaga tttccttggg aggggtgcct tcaatgtcat caacattgga gcatgggcta 120
gcaggccagt ccaggggtcc gcggtagatc ttagtcatgg actgcatctg gggctccatt 180
tgaagaacca tttgtagttt tacagcttcg attctggaag agacaaacgt aacaaggagg 240
ttaagatac aaggattgaa atgtacggcc tgaagtgcag gggcatatga gtgtgggcgg 300
tgcaagtggg gtttccttta gaaaaactcc gatacaatag ggcataata tttctaggaa 360
gccacattct ccatagaagc tctcggtaag gggagctact ggtagtacag cagcatacag 420
ggggtgcagt gagagtgaag gggggtaaga gaacagtaaa aagaaaaata tgacaaggga 480
gggccaagag gatctacgat tctagttact ttctcacgg ttgtcgctg aagagcaggc 540
gcagatcctc tagaggttca caggaatagc tagcattgtc tgctggattt tcgggttcct 600
ttggcagtat ccagggtttg gctcagagtgt gacttatcca agactccact ccagccactt 660
aactgcgggt agggtagata aaatgactgg gtagggctct tcccaggatg tgtgtaggga 720
tggggaatta aagggaagg gacttgacta ataccatgtc accagggtgg aataattcct 780
ttccctcctc tcagggacag gttccctgta atgttttaag aactcgttga tatttggcta 840
aggaggtgat gtctgcaact aagttggcgg tctctcagtc aagcacaagg tcattggtta 900
ggaagggtcg tccatacagc atctcatatg gactaagtcg tgctttttgg ggacagtttc 960
ggattcttag taaggctata ggcaacagag caggccatgc aagggtgggtt tcttgggtta 1020
gcttttttag atgtcgtttg agtgtttcat tcaattttct aacttttctt gaggatcgtg 1080
gcctccaggc acagtgtgaa tgatattgta tacctaacgc ctgggatact ccctgcgtta 1140
ctgcagcctt gaaattgggg ccattgtcac tctgtaaaac tcagggaagt ccgaatctgg 1200
gaattatttc atgaattagt acttttatta cctcttgggc cttttctgtc ctacaaggga 1260
aggcctccac ccaaccagtg aaagtaccca gattagtaga tactgaaatc tctgagattt 1320
gggcatgtgg gtaaaatcta gttgctagtc ttctcctggg taatggcctg ttctttgttc 1380
tcctgaagga gcttggcaat aaggcagggg attatttctt tggcacactt cacaggccct 1440
gactatctgc ttgacagttt tgaaaaggcc tgggtccagta aataatgatt tggccatctg 1500
atgggtgctg tcaatgccta agtgaaaggg ctggtgaagg gtttttaagta atttccattg 1560
gttagctgca ggcaaaagta ttttttcttt ggtggctggc catcctgagg agaggaaact 1620
atgtcctcgt gagtttcccc attccatttc ttctgctgag tactggagct tgggtttcca 1680
gaggggatta cccatacta ggggtccttc tgtaagcatt tctaattggag agtccctgct 1740

```

<210> 62

<211> 7140

<212> ADN

<213> Homo sapiens

<400> 62

```

ttggtcttaa gaacacaaat gatatggctc caatgactgg aggaacacca gggtccttgg 60
tctcacgctg atttagataa aacgactgtc aggcctctga gccaagcta agccatcctc 120
ccctgtgacc tgcacgtata catccagatg gcctgaagta accaaagaat cacaaaagca 180
gtgaaaatgg cctgttcctg ccttaactga tgacattcca ccatttgtat ttgttcctgc 240
cccatcttaa ctgagcgatt aacotttgtg aattccttct cctggtcaa aacctcccc 300
actgagcacc ttgtgacccc cggccctgcc cctaagagaa aacccccctt gattataatt 360
ttccactacc caccaaaatc ctataaaatg gcccacccc tatctccctt cgctgactcc 420
tttttcggac tcagcccgcc tgcacccagg tgaaataaac agccttggtg ctacacaaa 480
gcctgttttg tggactctct tcacacggac gctcatgaca tttggtgcca aaacctggga 540
taggaggact ccttcaggag accagtcccc tgtccttgcc ctcaactctg gaggacatcc 600
acctacaacc ttgggtcctc agaccaacca gccaaggaa cagctcacca atttcaaate 660
aggtaagcag tcttttctct ctcttctcca gctctcttgc ctacccttca aactccctct 720
ctcactacc ttcaattctc ctgtccttcc aattccagtt ctttttctat tctagtagag 780
acaaaggaga cacattttat ccatggaccc aaaactccag caccagtcac ggacttggga 840
agacagtctt cccttggtgt ttaatcactg cggggacgcc tgcctgatta ttcaccaca 900
ctccattggg gtctgatcac ggtggggaca cctgccttgg tcaactaccc acattccctt 960
ggtggtacgt caactgcaaa agcaggggac gcctgctttg gctgctcacc caccctcttc 1020
tctgtgtctc tacctttctc tttaaactta cctccttcac tatgggcaaa cttctgccct 1080
ccattcccc ttcttctccc ttagcctgtg ttcttaaaaa cctaaaaact cttcaactca 1140

```

cacctgacct	aaaaccta	aa	tgcccttattt	tcttctgcaa	cactgcgtgg	ctgcagtaca	1200
aacttgataa	tagcttttaa	aa	tggccagaat	atggcacttt	caattttctcc	atcctacaag	1260
atctagataa	tttttgtgga	aaaatggaaa	aatgggtctga	gatgcctgac	gtccaggcat		1320
tcttttacac	attgggtccct	ccctagtctc	tgctcccaat	gcgactcatc	ccaaatcttt		1380
cttctttctc	tcctgtctgt	tccttcagtc	tcaccccaa	gctctgagtc	ctttgaatcc		1440
tcctttgcta	cagacccatc	tgaactctcc	cctcctcccc	aggctgctcc	tcaccaggcc		1500
gagccaggtc	ccaattcttc	ctcagcctct	gctcccccac	cctataatcc	ttttatcacc		1560
tcctctcctc	acactcagtc	oggttacag	tttctgtctg	tgactagccc	tccccatct		1620
gcccacaat	ttcctcttaa	agaggtggct	ggagctaaag	gcatagtcaa	ggttaatgct		1680
cctttttctt	tacttgacct	ctcccaaate	agttagcggt	tacgctcttt	ttcatcaaat		1740
ataaaaaccc	agccagttca	tggcccatct	ggcaacaacc	cttacaggct	ttacagccct		1800
agaccctgaa	gggtcagaag	gocgtcttat	tctcaatatg	cattttatta	cccaatccgc		1860
tcccaacatt	aaataaagct	ccaaaaatta	aattctggcc	ctcaaaccac	acaacaggac		1920
ttaattaacc	tcacttcaag	gtgtacaaga	atagagtaga	ggcagccaag	tagcaacgta		1980
tttgagttgc	aattccttgc	ctcaactctg	agagaaaacc	cagccacatc	tccagcaaac		2040
aagaacttca	aaacacctga	actgcagcag	ccaggcggtc	ctccaggacc	acctccccc		2100
ggatcttgct	tcaagtgcgc	gaaatctgac	cattggggcca	aggaatgcct	gtcggccagg		2160
attcctccta	agccacgtcc	catttctgca	ggaccccaact	ggaaatcgga	ctgtcccaact		2220
cacccggcag	ccaatcccag	agcccttgga	actctggccc	aaggctctct	gactgactcc		2280
ttcccagatc	ttctcggtct	agcagctgaa	gactgacact	gcccgatcac	ttcagaagtc		2340
ccctggacca	tcacggatac	tgagcttcag	gtaactctca	cagtggaggc	taagtccatc		2400
ccctgtttta	tcgatacagg	ggctaccac	tccacatcac	cttcttttca	agggcctggt		2460
tccctttccc	ccataactgt	tgtgggtatt	gacggccaag	cttcaaaacc	ccttaaaact		2520
ccccactctc	ggtgccaaact	tggacaacat	tcttttatgc	actctttttc	agttatcctc		2580
acctgcccag	ttcccttatt	aggccgagac	attttaacca	aattatctgc	ttcccgcact		2640
attcctgggc	tacagccaca	tctccttgcc	gcccttcttc	ccaacccaaa	gocctccttc		2700
tatcttcctc	tcatatcccc	ccaccttaac	ccacaagtat	gggacacctc	tactccctcc		2760
ctggcaaccg	atcacacgcc	cattactatc	ccataaaac	ctaataccac	ttaccctgct		2820
caatgccagt	atcccatacc	acaacaggct	ttaaagggat	tgaagcctgt	tatcacttgc		2880
ctgctacagc	acgggcttct	aaaacctata	aactctccat	acaattcccc	cattttacct		2940
gtctaaaaac	cagataagtc	ttacagggtta	gttcagaatc	tgcaccttat	caaccaaatt		3000
gttttgcccta	tccaccctgt	agcaccaca	tctgtaactc	ttttgtcctc	aatgccttcc		3060
cccacaactc	actattccgt	tcttgatctt	aaagatgctt	ttttcactat	tccctgacac		3120
ccctcatccc	agcctctctt	tgtctttacc	tggactgacc	ctgacacca	ctagtcacac		3180
cagcttacct	gggctgtact	gccgcaaggc	ttcagggaca	gccctcatta	cttcagccaa		3240
gctctttctc	atgatttact	ttctttccac	ctctctgctt	ctcaccttat	tcaatatatt		3300
gatgaccttc	tactttgtag	ccctcctttt	aaatcttctc	aacaagacac	cctcctgctc		3360
cttcaacatt	tgttctccaa	aggatatcgg	gtatccccct	ccaaagctca	aatttcttct		3420
ccatctgtta	catacctcgg	cataattctt	catgaaaaca	catgtgctct	ccctgccaat		3480
tgcgtctcca	actgatctct	caaataccaa	cctcttctac	aaaacaacaa	ctcttttccc		3540
tcctaggcat	ggttggtatac	ttttgccttt	ggatacctgg	ttttgccatc	ctaacaaaaat		3600
cattatataa	actcacaaaa	ggaaacctag	ctgaccccat	agatttctaa	tcctttcccc		3660
actcctcttt	ccattccttg	aagacagctt	tagagactgc	tcccacacta	gctctccctg		3720
tctcatccca	acccttttca	ttacacacag	ccgaagtgca	gggctgtgca	gtcgggaattc		3780
ttacacaagg	accgggacca	tgcctgttag	cctttttgtc	caaacaactt	gaccttactg		3840
ttttaggctc	gccatcatgt	ctccatgcgg	tagcttccgc	tgcctaata	cttttagagg		3900
ccctcaaaat	cacaaactat	gctcaactca	ctctctacag	ctctcacaa	ttccaaaatc		3960
tattttcttt	ctcacacctg	acgcatatac	tttctgctcc	ccggctcctt	cagctgtatt		4020
cactctttgt	tgagtctccc	acaattacca	ttcttccctg	cccagacttc	aatctggcct		4080
cccacattat	tctggatacc	acacctgacc	ctgatgattg	tatgtctctg	atctacctga		4140
cattcacccc	atttccccat	atttcttctc	tttctgttcc	tcatgttgat	cacatttggt		4200
ttactgacgg	cagttccacc	aggcctgacc	gccactcacc	agcaaaaggca	ggctatgcta		4260
tagaatcttc	cacatccatc	attgaggcta	ctgctctgcc	ccccccact	acctctcagc		4320
aagccgaact	gattgcctta	actcgggcct	tcactcttgc	aaagggacta	cacgtcaata		4380
tttatactga	ctctaaatat	gccttccata	tcttgcacca	ccatgctggt	atatgggctg		4440
aaagagggtt	cctcactacg	caagggtcct	ccatcattaa	tgcctcttta	ataaaaaactc		4500
ttctcaaggc	tgttttactt	ccaaaggaag	ctggagtcc	acactgcaag	ggccacccaaa		4560
aggcgtcaga	tcccattact	ctaggaatg	cttatgctga	taaggtagct	aaagaagcac		4620

ctagcggttcc	aactttctgtc	cctcatggcc	agttttttctc	cttcccatca	gtcattccca	4680
cctactcccc	cattgaaact	tccgcctatc	aatctcttct	cacacaaggc	aaatgggttct	4740
tagaccaagg	aaaatatctc	cttccagcct	cacaggccca	ttctattctg	tcatcatttc	4800
ataacctctt	ccatgtaggt	tacaagccac	tagtccacct	cttagaacct	ctcatttctt	4860
ttcatcggtg	aaacatatcc	tcaaggaaat	cacttctcag	tgttccatct	gctattctac	4920
tacccctcag	ggattgttca	ggccccctcc	cctccctaca	catcaagctc	ggggatttgc	4980
ccctgcccag	gactggcaaa	ttgactttac	tcacatgccc	tgagtcagga	aactaaaata	5040
cctcttggtc	tgggtagaca	ctgtcactgg	atgggtagag	gcctttccca	caggggtctga	5100
gaaggccact	gcagtcattt	cttcccttct	gtcagacata	attccttggg	ttggccttcc	5160
cacctctata	cagtccaata	acggagcagc	ctttatttagt	caaatcacct	gagcagtttt	5220
tcaggctctt	ggtattcagt	ggaaccttcg	taccccttac	tgctctcaat	cttcaggaaa	5280
ggtagaatgg	actaatggtc	ttttaaaaaac	acaccccacc	aaactcagcc	tccaacttaa	5340
aaaggaggat	agagcccaaa	aactcgcaac	caagctagta	attatgctga	acccccttgg	5400
gcactctcta	attggatgtc	ttaggctctc	ccaaatctta	gtcctttaat	atctgttttt	5460
ctccttctct	tattcggacc	ttgtgtcttc	cgttttagttt	ttcaattcat	acaaaaccgc	5520
atccaggcca	tcaccaatcg	ttctatacaa	taaatgctcc	ttctaacaac	cccacaatat	5580
cgccccctac	cacaaaatct	tccttcagct	taatctctcc	cactctaggt	tcccatgcgc	5640
cccataatcc	ctctogaagc	agccctgaga	aacatagccc	attatctctc	cataccaccc	5700
ccaaaatttt	tgtgccccca	acacttcaac	actattttac	attatttttc	ttattaatat	5760
aagaagacag	caatgtcagg	cctctgagcc	caagccatca	tatccccctgt	gacctgcaca	5820
tatacatcca	gatggcctga	agtaactgaa	gaatcacaaa	agaagtgaag	atggcctggt	5880
cctgccttaa	ccgatgacat	tcacccactg	tgatttggtc	ctgccccacc	ttaactgagc	5940
aattaacctt	gggaaattcc	ttctcctggc	tcaaaacctc	ccccactgag	caccttgtga	6000
cccctgcccc	tcactacccc	acccaaatcc	tataaaatgg	ccccacccca	tctcccttag	6060
ctgactcctt	ttttggactc	agcccgccctg	caccaggtg	aaataaacag	ccttgttgct	6120
cacacaaagc	ctgtttgggtg	gactctcttc	acagggacgg	gggtgacaac	aacacggaca	6180
cacatggagt	ggttttaagg	agcagagagt	ttaatacgca	aaaaagaagg	aagaggctcc	6240
cctgtacaga	cacagagga	gggggctcca	agccgagaga	aggaaacccc	atgtgcagtg	6300
gaaaagtgg	tgattatact	gggaggtgg	aggaggcgg	gtctgatttg	cacagggcc	6360
aggggattgg	gttgaccagg	tgtatcattc	atgtaccccg	caaaaaacct	ggccctccca	6420
cctcagccct	ttaatatgca	aatgtgggtt	gccatgatgt	tctgaaaaca	catgaattat	6480
ctggaggggg	ccatgacact	tggtacatgt	gctgacaaga	agaggggtgg	aatcgccatg	6540
gtggccatgt	tgggtggacc	tagtttttaa	tagcctgcat	ttgcataatca	aagtttgctg	6600
gcctggctct	ttaaactgtc	ttttctgtta	gaaaagggaat	ggtttggaat	gggtgaggtg	6660
tgcttcttat	tacaagaaaa	tttccaaaaa	cctttactct	ttctagctgc	caaaaaacta	6720
tttcttaata	acttatgtat	taccataatt	aggcagcacc	aaagatccct	gcaggtcaga	6780
ccactgcaat	taacatgctg	gctttactgc	tgattatgg	agctgcatcc	acctagcctc	6840
tcatattgca	actgcctgac	ctctgccacc	ccacgagcca	cttatcccca	cttataatca	6900
gcccatttctg	attgtaacat	ctgccactta	ttcccagcgt	tgtggtatat	cctatagatg	6960
aattcattca	acatccattc	caacaccacc	tctcttgcc	tcctatactc	tctggagagt	7020
gaattactga	gtcacatgat	cttcactgca	gtcatttgtg	gctatgtgac	atagttctgg	7080
acagtgaaca	tagacagaag	tccttggggc	gggcttctct	tctgggatga	gggcaaaacg	7140

<210> 63

<211> 44100

<212> ADN

<213> Homo sapiens

<400> 63

tgccctttatt	tccgtaggct	ggatcatatgg	cgctagcact	cacataaagc	taccgaggag	60
agcgaatgaa	acaaaaatca	ctttaccttc	acagcacgag	gccgtcgtcc	ctctcgatat	120
ttggcccggtg	tgctgcatac	cgccctctgg	acgtgggtgat	caaataaact	ccctagctcc	180
ccgcccgtctg	acgccatctt	gcctactttg	atcctcgcag	ggaggacaac	atccgcccta	240
ctgagctccc	ttttatccaa	taagagagcg	ggatgagtta	aggagtgcc	ggattggctg	300
gagaatcgac	agcgtcggcc	atcgtttcct	gcgtgcgaag	atttgatgaa	cgagggtgccg	360
cccccgagcg	gctcggcgga	gaggcgcggt	gggtgacaga	agctttcttg	tcccaccac	420
tacaggctta	cggcaggatg	cgcagcgggg	agagggggcg	gggccgcagg	gggcggggcc	480

gatcgatctc	ctccggetcc	gacgtcctcg	gcctgcgcgg	tcccgggtcc	tttgcgggcgc	540
taggggtgggc	gaacccagag	cgacgctccg	ggacgatgtg	gggcagcgat	cgccctggcgg	600
gtgctggggg	aggcggggcg	gcagtgactg	tggccttcac	caacgctcgc	gactgcttcc	660
tccacctgcc	gcggcgctctc	gtggcccagc	tgcattctgct	gcaggtaacc	tgcgggcccc	720
gagccacctg	atcttcagcc	tggggctcgga	cgaggccgaa	gcctctcagg	gacgcggcgg	780
gacaccggct	gccaccggg	cgccgcgaa	gcgcgcagag	atcagggtcc	ctcgacggca	840
gggcccctct	gggtagtctc	tggatccac	aagtcagtg	cagccctggg	ctcgtcttat	900
cccaggtctt	ttcacttggt	gaaactgaac	ctagaaacgt	cctaataattc	taccactgtt	960
tttataaata	ttccttattc	caggctggaa	aagctcctga	gaagtgggtt	gtttttatta	1020
ttttaaaagg	tgttttcctt	gccagccatt	tccagttaac	ctgcgctgct	gccgtccggg	1080
ccgcgagagc	gggacgcaga	gttggtggcg	gagccctgt	cggttcccgg	ggactaagca	1140
ccgcgtccca	tgagcgggaa	aggttaatac	aatgatggtt	ctgccctgcg	tcgctgacgc	1200
ggaacacagc	tgtagtgtgt	taggaacaca	taacgtagt	aagatcactt	gaagctctgc	1260
gatcagtcgc	ccttctggac	gttggtggtt	ggatgtttca	cagttctaac	cactgggtgga	1320
gatacagcgt	ccatattttc	ataattaaaa	atagaggcac	atggtctcac	gagtttgagt	1380
gtacttatgg	gggcaaaagg	acggcgtatt	tgaatccctc	ataaatcctg	gatgcattgt	1440
accaccaggt	ggctaattct	tgcaatgaat	agagtttgca	ataatttcaa	gcattccctc	1500
tttccacttg	agttacttcc	ccatacctag	gggaagatat	ttttgggtcca	ctgaaaacat	1560
gagttcagca	gaatcctcct	atcatcgtcg	ttattatttt	ttaccactaa	gtagacaatc	1620
ttttggtttt	tgatgggctt	tatggctaga	gacaaatcag	tcactgtcac	caagttccag	1680
gtagaagttg	gttcagtgct	ctgtcagctt	cgatgggatt	tttcaacatg	ttttcaaatc	1740
tgcacttaat	agtaggaatg	ctttcttaca	gtaactctaa	tttgatccta	agatgtagt	1800
gttaccttac	attcatcact	gtttaagaat	ttagtggctc	tgatctttgt	tttaaatatt	1860
gagccttcgg	gaagtactta	taagaattaa	ttcatgcata	tctttttgaa	atgtaaatgt	1920
ctttagccct	ggaacaaatt	gctgtttctg	ttcagcccat	attagcagaa	taggtcaact	1980
ttactttcta	attatcaatg	taataagttt	attactttat	agattccata	aattctataca	2040
tttattcctc	gatgaattat	ataaatttat	agaatttatg	ttttatagaa	aattttgaaa	2100
gcatggaaaa	ttattaacaa	gaaaataagt	taccataat	cccagaactt	agagggtgact	2160
aatggtgaca	gtttggatca	aatcttccag	ttttgtttct	aatctttatt	tttaacataa	2220
atgaggtcct	gtatacacac	gtacagtttt	gtgtcctggt	gtttttattt	aatgttatta	2280
tgagtgtttt	atttttgttaa	aaggctcatca	ttttaagttg	ttaattagta	ttctagcaca	2340
aatttgccat	aattttattta	attgtttact	atgattgacc	atttagattg	tacttaattt	2400
ttaggcatta	gaagtgataa	actatatttt	aatcagacgt	tgaaaataac	acatctttgt	2460
ttagaaaaca	tcattttatt	tctggttgct	taggatagat	tcccagaatt	cttgggttag	2520
aggccataga	taattatgaa	agcagaaaga	ttcacaagtt	gggagttaat	acttgaattc	2580
ctttatttgg	ggtgaagcat	tgagtgcata	atacagatca	tgcagtaatg	ggaagaagg	2640
ttggaacaat	ggttttctgg	cctatgtcag	acttaccttg	aagcttttaa	gaatacagat	2700
gttctgatca	accctcagac	ctattaaatc	agacctaaaa	tcttagggaa	taggcttttag	2760
gcatctctaa	ttttaaaaaa	tttattcagg	ctacttggt	gcacaaaaga	gttgagacct	2820
actgtcctag	aatcatagaa	ttttaatgac	gatagagacc	ttaagcatct	aggctggttc	2880
tgtactttta	catgtaagga	aactggcatt	cctaggccag	taccattgcc	atgcagctaa	2940
tttgccctct	tgtctatagc	tcactctgca	tcacccaacc	taccgttctc	actgtttctt	3000
ctataacca	tctccttccc	acttctgttc	tcttactcat	gccattcttc	cctcagtcac	3060
ttttcttcc	tccatacaaa	ttccatgtct	ttaaaaagga	ataatcctac	ctcctccaca	3120
tagctttcca	attctctgtt	gcccacattt	gtctcccttt	caatacttct	ctgttggtgt	3180
atgtgacaca	tcacatttga	tatactctgt	actgtgtttc	aagtattgta	ttctcttgtt	3240
tactcaagtc	attatttcag	gactgactac	ccagttagatg	ctttaagtca	ggattttctca	3300
accttggcac	tgttgacatt	ttgagctgga	taattttttg	ttttgggggc	tctcctgtac	3360
attttaagat	gtttaacagc	acccttggcc	tctatccagt	agacgcctgt	actgcctccc	3420
cctatctgtg	acaaccaaaa	aggtcttcag	acattgtcag	atgtctactg	aaggacaaaa	3480
tcacctctgg	ttgagaacca	ccgcttcaac	taagttatct	tctctgtact	cagaacttga	3540
tgtgattgca	gcagggggag	aggattcata	tacacagtga	atgcaaacga	acctaaatca	3600
ccattcggat	attggccacac	aattttcatt	tcccttgtgt	tagcaagaga	taccctaggc	3660
tttggaacctg	attattccta	aggcattctg	atgtatggtt	ttacctgcag	atttccctgg	3720
aatactgata	cctcagtttg	gggtcaaagaa	gggtcaattaa	ttgattgatt	tgatttgact	3780
cctggaaaag	acgctccttt	ctagctgtct	ctttcttctc	tttacctgaa	tagccagggc	3840
tctgtggttc	aagtgaagta	ttttgacata	aaaattaaact	tagaacattg	gtctgcagag	3900
tttgctcaat	ataactgagc	acatattgtg	gctttatgga	gctggttact	actttttgac	3960

caaataaata	attagaagta	tttttcctcc	tcaataaggt	tcattttttcc	ttttttcagt	4020
gagctggtag	agtttccctt	tttgatattt	cagggcatct	ttcataatttc	catctcttaa	4080
gtttcttcat	atgaagtaga	atztatctgg	attatgtatt	gctgactctg	atgaaaaccc	4140
atagaaagca	tctggggcct	gatcaccttc	attcttgtaa	tagctcacac	ggttacagct	4200
gatatggtaa	cttaagactt	ttgattccaa	atctaggcaa	aatacactca	gttgaaagaa	4260
tttgtcagcc	agaacagttg	gactgttctg	tgaaaattgt	gagaaaaatt	acacaactaa	4320
gtgatacatg	atgatggcct	tcttaaatat	aaaattgtaa	taacatgggt	aattttccagt	4380
acgttatatt	gtcccagaag	tggtcccaac	attgtttgaa	atttgtctca	tttaaagaaa	4440
cataagctgg	ctatggtggc	tcacgcctgt	aatcccagca	ctttggggagg	ctgaggcagg	4500
cagatcacct	gaggtcagga	gttcgagacc	agcctggcca	acatggtaaa	accccatctc	4560
tactaaaaat	acaaaaatta	gccgggcatt	tggtggggggc	ctgtaatccc	agctacttgg	4620
gaggctgagg	caggagaatt	gcttgaatct	gggagggtgga	ggttgcagtg	agccgagatt	4680
gtgccactgc	cctccagcct	gggtgacaga	gtgagctctc	gtctcaagaa	aaaaaaaaaa	4740
aaaagcaaga	aacataaaga	ctgggcatgt	tggtctcatgc	ctgtaatccc	agcactttga	4800
gagactgagg	tggaagatc	acttgagccc	aggagggttaa	ggctgcagtg	agccgtgatt	4860
ttgccactgt	actcgagcct	gggcaacaca	gtgagatcct	gtctcaggaa	aaaaaaaaatt	4920
gcatgtaaat	gaatgaattt	gatattttaat	atttttaaatt	atgaaaactg	ttctgtagag	4980
atgtagatct	tgccatgttg	cccaggctgg	ctttgaactt	ctgggctcaa	acaatcctcc	5040
tgtctcagtc	tcccaaagta	taaagattac	acatgtgagc	cactgcacct	ggcctaatat	5100
ttttaactta	atgaatttat	tttgatataa	ataaattaat	aacactgaag	cttctctgata	5160
taataagtct	ttttgtgtgt	gtgacgggtt	ctcactctgt	tgcccagact	ggagtgtaat	5220
ggcactatca	tggtcactg	tagcctcaac	ctccctgact	caagtgatcc	tcccacctcg	5280
gcttccctgag	tagatgggac	cacaggcgta	tgccaccaca	cctggctgat	ttttaaaaat	5340
tattattgat	acataattaat	aaaattattt	ttatttttaa	aatgatatat	gtggctgggc	5400
atggtggctc	atgcctgtaa	tcccgacagt	ttgggaggcc	gaggtgggag	gctcacttga	5460
gaccaggagc	ttaagaccag	cctaagcaac	atagtggat	cccactctcta	tagaaaaaaa	5520
aaatggctag	gtgtgggtgg	gtatgcctat	attcccagct	actcaggaga	ctgagggtgag	5580
aggattgcta	gagcccagga	gtttcaagtt	acagtgcact	atgattgtgc	cagtgcactc	5640
cagcctgggc	aacagagcaa	aatcctgtct	caaaaaaaaa	aaaagtccga	aaatgcttat	5700
gatgaatat	aagtatgga	aaaggatatt	aaattgtgcc	tatatgaaca	caactatatg	5760
aaaaacttgc	acatagagaa	aaggattaac	aagaaataga	ccaaattggt	cacatgggtg	5820
tcttgtttgt	ggagagaata	tcagtagttc	atltgtttcc	ttccaagtgt	atatgttttc	5880
cgagggtctct	ataatgagtt	tgtaattggt	taatcataga	aaaccctttt	ttggctcctg	5940
gccacaaact	tacatgtttt	aatgtaattg	cttttttaat	gagaataaat	gttatatttt	6000
gcttttttaa	aacctatatt	cccatagtta	tatgagccct	tacaattatt	aagaggctgc	6060
ataatataac	gtttctggaa	gggtacagaa	gaacacgcag	taattacctc	tgagaacaga	6120
gacatggctt	cacattttac	ccttttgtac	gttttgtgct	tttgccacat	gcattttatta	6180
ttcttccaat	aaataagtaa	ataaatatgg	attgtatact	ccatctgggt	ggtgtttcat	6240
aattctaaaa	ttatattgct	acatttttaa	agatgatatg	tgtttctact	tattaacgta	6300
tatgttaaaa	tagtaaattt	atatcttatt	taataatttc	cctattgata	gacattttaag	6360
acagtctcaa	gtgttcacta	tcatagaaaa	tactgcacag	atagcttttg	ctatagtttc	6420
ttttttcttt	gaatcgttaa	ttgggaataa	atgctcaaat	agttatatgt	ggctcaactg	6480
ctattttaagt	ttattgactg	actgctgcca	ttttgaattc	tgaaggggtt	gattaaattt	6540
ataatgctgc	cataagaata	taagggtatt	ggcttcatta	gcattccacca	gcattgggtg	6600
ttggaaatga	ttatagattt	ttaaatgcta	caacaaatgt	agataacaga	gaactatcta	6660
tagaactctt	tttggacatg	tgaattgtaa	taatagttaa	ttttcatgtg	aatccagaaa	6720
aatgtatacg	aaaacctttt	ttcctctcat	ttcttatatg	aatagaatca	agctatagaa	6780
gtggtctgga	gtcaccagcc	tgcatctctg	agctgggtgg	aaggcaggca	ttttagtgat	6840
gggggacagg	taagcacatg	tgatggcaat	aactttcttc	taatatcaca	taatatagca	6900
atagaaataa	aattaaaagt	ttagattttt	tgtaaaggga	ggtgagatgt	cacctaatat	6960
gtatgctatt	atgtaactag	tctaggatat	tgaagctgac	tatactctgt	ttttaggtca	7020
ttatcttgta	gtttaccata	ctccctactt	gcttcttatt	ctactattta	actcattttc	7080
cacatccctt	aattttggtt	tcatgaaatt	atttttcctt	ctgaattact	aggttctact	7140
tactattatt	aaactttatt	tctgacatat	tttataacct	tccatgggtc	cacttgatta	7200
aaaataaaaa	attcagctgg	gtgcgggtggc	tcacacctat	aatcccagca	ctttggggagg	7260
ccaagggtggg	cggataattt	gaggtcagga	gttgagagacc	agcctgcccc	acgtgggtgaa	7320
accccccttc	tctactaaaa	attcaaaaat	tagctgggca	tggtggcagg	tgctgtaat	7380
cccagctact	caggaggctg	aggcaggaga	attgcttgaa	cctgggagggt	ggagggttgca	7440

gtgagctgag	attgcactgc	tgcacttcag	ctgggtgaca	agagcgaaac	aatgtcttga	7500
aaaaaaataa	aaaataaaaa	attctacaac	acaggggttat	tattttttcca	tttttgtttt	7560
cccttatgag	tttaatatgt	ttagattata	aacctgaaag	cttgaatacc	tatgtctatc	7620
ttttgttttc	ttatgtttat	caagttattc	ctttaaacat	tttctaact	gtaagaataa	7680
tgtgaggctg	ggctcaatgg	cttatgcctg	taatcccagt	gctttgggag	gccaaaggtg	7740
gaggaccact	tgaggccacg	agttcaagat	tagcctggct	aggcaacata	gcaagaccct	7800
atctctataa	aaaaattaaa	aaaatttagct	gggcatggta	gcaaatgctt	gtagtcccag	7860
ctactcagca	gactgaggta	ggaggaatgc	ttgagaccag	gaatttgagt	gacctatgat	7920
tatgcactcc	agcccgggca	atagcaagac	cctatctctt	aaaagaagaa	gatgtagtaa	7980
taatacatat	tcattataac	tattttacca	ttgaaagtaa	aaaatgagtt	tttacctttt	8040
cccagtccca	tcctcagaat	gggatctctc	gtagaccttt	aggattggaa	gaatgagatc	8100
attcatattt	tctgcaatta	ttaccccaca	aaatatttca	gatacctttc	catgtattac	8160
aaacaatgtg	catttaacat	gtctctctct	ttctctctct	ctctgtgtgc	gtcttcatga	8220
tcctctgttg	cagccctgcc	agtaagacac	tatctcctga	agaatcactg	ataggaacag	8280
aaagtggact	ggctaggcca	ggagtcctta	gcttcttagg	gggcaggagc	tgctttgtgc	8340
tttctcagaa	tcagatata	atgtggactg	aaacatttaa	aaacagaata	gccaaaggtg	8400
ctatacgttt	aaaacttata	tagatggggc	tacattgctc	tctattacta	atttcccattg	8460
acaatacacg	agagtggcat	gtctttttaa	cttggtttga	gcacagacta	atcttgttta	8520
tgcatgtttt	ttgatgagaa	taggctactc	atgagaaaac	tgtaaaccta	acactagtcc	8580
cttgcatact	ctaaaattgtt	gctagaatct	taaaatttta	gcaccagacg	gaccttagaa	8640
atcattaact	ttggtgcttt	gttctacaat	acaaggagat	ggaatatatt	accaggattt	8700
gcttagcagg	ttacagttct	gccctctgag	taccagcac	ttccctgtgg	gcaacatcaa	8760
cttcctgatt	ttcaagtctt	aattagtact	ctgaagaatc	ctacttgttt	ttaaactcca	8820
tttgctttga	agtgacttta	cctgattttt	ttagatccct	tattgcagca	atgccactaa	8880
gaaactgagt	ctctagcttc	ttggtgggca	ggagctgctt	tgtgcttgct	cagaatcatc	8940
cttttcagta	agggagatat	tgaagagaaa	tctactgagg	agtctggggg	tgaggcactc	9000
agggaaatcc	tgctccagtc	cacaaaagca	gagagggaag	gttggttacc	tagagtattt	9060
aacatgcaga	ggctttggat	tttactcctt	taatccttgg	aatgcctat	ggaaggggaa	9120
aggaagtaag	atggtgactc	cagcttatag	acatactagt	gttacatata	tttaaactat	9180
aataggaggg	tattattagt	tttacttaac	tttcaactgt	gaaggattat	acttctcaat	9240
atttgtctcc	agtgtctatt	tcagtgtatt	tttcactttt	cttgaagcag	catgtctgtt	9300
gcaaaacttc	tagaaataat	gagaatat	atatattaga	tcaagccata	acttgatgat	9360
atagtcat	cttcttatat	tttttactta	cattttttaca	ttttaatgat	tactttcatt	9420
tttgaaaaac	atgtcatgct	gagatgtatt	tttcttcatt	ctgtaattag	ttatgaaaca	9480
gtttttccta	aaatgctgag	tatatcaagt	cttggtctaa	aataagtaat	aaatatttgc	9540
cacatgaaag	actacacata	tagccagggt	cagtggcttg	cacctgtttt	ccagctacc	9600
caggaggctg	aggcaggagg	attgcttgag	cccagggttt	ccaggctgca	gtgaactatg	9660
attgtaccac	tctactccag	aatgggtgac	agagccaggc	cccatctctc	aaaacagaaa	9720
agaaagatta	catagactac	atatacacc	ccatccaaaa	catacacaca	catctactta	9780
acctaaaatg	gtaagaagat	aacttcttat	tttctaatat	atgacacaga	aaagtttttt	9840
taaagttagt	ttaaattttt	aattttttct	aggatattct	caagccatgt	tccatgttgg	9900
tatcttgtca	acaagttgag	gtggaacccc	tctcagcaga	tgattgggag	atactggtaa	9960
agaaaaccaa	ataagaacta	tctcatttaa	ggttaaatta	cttcacaata	tcaatgtctt	10020
tagctttctc	taagctttat	tatatattct	gagttgggtt	tgaattataa	gaatgaattg	10080
gggccaggca	cagtagctca	tgcttatagt	cccagcactt	tgggaggcca	aggcagggtg	10140
attgcttgag	tccaggagtt	caagaccagg	ctgggcaaca	tggtgaaacc	ccgtatctac	10200
taaaaataca	aaaattagcc	aggcatggta	gtgcatgcca	ttagtcccag	tcacttggga	10260
ggctgaggca	ggagaatcgc	ttgagcccg	aaagtcaagg	ctgcagtgag	tcaggatctt	10320
gccattgtac	tccagtctgg	aaaacagagt	gagaccttgt	ctcaaataaa	aaaagaatga	10380
attgatagag	atctaagtga	caacctgaca	actataggta	ataaaattgt	attggggatt	10440
catgttaaat	gagtagattt	taactactct	taccacaaaa	acacaaaagt	gggtaactgt	10500
gagatgatgt	atatgttaat	ttacttcact	atagtaacca	ttatactatc	tatatgtagc	10560
tcataacacc	atgtcgtgta	tattaaatat	gcacattaaa	atttgttttt	taaaaaaaga	10620
attgagattt	tttttaacta	gatatggagt	ggacaaaaatg	taaagtgaat	tgatcttttc	10680
gtctgttggt	tctaggagct	gcatgctgtt	tcccttgaac	aacatcttct	agatcaaatt	10740
cgaatagttt	ttccaaaagc	catttttctc	gtttgggttg	atcaacaaac	gtacatattt	10800
atccaaattg	gtaggtgcta	ttgtaatat	tgctgtcata	ttctacacta	tagcattgag	10860
tccaaagtag	aaatgaatgt	gcactaatga	gctttatttt	ctacacagtt	gcactaatat	10920

cagctgcctc	ttatggaagg	ctggaaactg	acaccaaact	ccttattcag	ccaaagacac	10980
gccgagccaa	agagaataca	ttttcaaaag	ctgatgctga	atataaaaaa	cttcatagtt	11040
atggaagaga	ccagaaaagga	atgatgaaag	aacttcaaac	caagcaactt	cagtcaaata	11100
ctgtgggaat	cactgaatct	aatgaaaacg	agtcagagat	tccagttgac	tcatacatcag	11160
tagcaagttt	atggactatg	ataggaagca	ttttttcctt	tcaatctgag	aagaaacaag	11220
agacatcttg	gggtttaact	gaaatcaatg	cattcaaaaa	tatgcagtca	aaggttggtc	11280
ctctagacaa	tatttttcaga	gtatgcaaat	ctcaacctcc	tagtatatat	aacgcgtcag	11340
caacctctgt	ttttcataaaa	cactgtgccca	ttcatgtatt	tccatgggac	caggaatatt	11400
ttgatgtaga	gccagcgttt	actgtgacat	atggaaaagct	agttaagcta	ctttctccaa	11460
agcaacagca	aagtaaaaca	aaacaaaatg	tggtatcacc	tgaaaaagag	aagcagatgt	11520
cagagccact	agatcaaaaa	aaaattaggt	cagatcataa	tgaagaagat	gagaaggcct	11580
gtgtgctaca	agtagtctgg	aatggacttg	aagaattgaa	caatgccatc	aaatatacca	11640
aaaatgtaga	agttctccat	cttgggaaag	tctgggttag	tataaatttt	ataacttggg	11700
agaaatttta	tgtggcttaa	acatccccaa	attatgaatt	agaatagtat	ttcatatata	11760
aattgaaaaat	caatttaaaaa	gaaacacagt	gcctaaaagg	acttggggga	cacattttacg	11820
ctttgcagta	aagtccttgt	ttggataaag	attgtatggt	ttctggccaa	gtaagcttga	11880
ataggtacaa	gcttagatag	gttcaggcca	gagaggtaaa	aattacttgc	ctgagattgc	11940
atagctagtg	ttacaactag	gattcaaaacc	caggcagatt	gacttggggg	ttcatcagga	12000
tggagtgcct	tacaaagcct	cccatcttta	atgcttgacg	atttggtccc	cagttaccga	12060
aagcaacttg	ttaatattag	ggaaaagggc	cagtgtaggg	agagatccat	ggcatgaggt	12120
aaccttcctg	ctgcatgtgg	tggcacctgg	attggaatgc	atccaggagc	tgcttaccct	12180
gccggtgtct	gctctttaat	ttgtgtataa	cggagaggaa	gtagacaggg	caactagtgc	12240
tccagccctt	catcctggcc	acaaatatta	atgctacctt	tatatgacat	aagtcactag	12300
tccattttatt	ggaacctaaa	tttgaaccac	tgtaaagtta	gacttcatag	tgataaagag	12360
aggaacttgt	taggaaaagag	aataaaaatag	aaagagaagg	ttgtctcctt	ttgtagattt	12420
tttttttttc	tccaacagtt	ttacctgtga	cctttatata	aataactgac	aaagcattaa	12480
tctctttggc	ctacatcatt	ttcttttcta	tttttttttt	ccacaagatg	gagtttcaact	12540
cttcttgccc	aagctggagt	gcagtggcct	gatctggctc	actgcaacct	ccgctcccca	12600
cgttcaagtg	gttctcctgc	ctcagcctcc	tgagttagctg	ggactacagg	cagtcaccac	12660
cacgcctggc	caattttttg	tattttttagt	agaaaactggg	tttcaccatg	ttagcagacc	12720
tggctctggaa	ctcctgacct	caggtgatct	gcctgcctcg	gcctcccaaa	gtgctgggat	12780
tacaggcatg	agccactgct	cctggccggc	ctacatcatt	ttctaaagct	ccagaccatt	12840
cttttctttt	cttttctttt	cttttctttt	cttttctttt	cttttctttt	cttttttctc	12900
ttctcttctc	ttctcttctc	ttctcttctc	ttctcttctc	ttttcttttc	tttttttaga	12960
ttagaagctt	gctttgttgc	ccaggctgga	gtgcagtggc	accacctcca	ctcactacaa	13020
cctccactgc	ccaggttcaa	atgattctcc	tgccctcagc	ttcagagtag	ctgggactac	13080
aagtgtgcgc	caccactcct	ggctaatttt	tgtattttta	gtagggacga	ggtttcacca	13140
tgttgccag	gctagtcttg	aactcctggg	ctcaagtgat	ccgcctgcct	cagtcctcca	13200
aggtgctggg	attacaggcg	tgagccactg	tgccctggcct	cagatcatta	ttttctgtta	13260
gctttaaact	gtccgttcag	gagatcccac	tgcatcctca	aattcaaaaat	atctaactac	13320
gagcttatga	tttagtgggt	tctgtcatta	gatgggaata	tcctttttatt	tccttgaaat	13380
tatatgggtga	gaacagggag	aagtgtctgat	ggtaaaagtcc	tgtgattaag	atagcaataa	13440
ggactccgcc	cttcccactc	cactgaaggt	tgaagagcca	tggacaatga	gaagtccacag	13500
taggtgaaat	caggtactaa	aatggacttg	gcttgagaga	tcaaaaattga	tcacttggtg	13560
atacaactaa	caaattcatg	ttaaactgaa	cctttattac	cctgtgaagc	atgggtgatta	13620
aaaaaaaaca	acaaacaaaac	aggaaaacttg	attgtttaa	tctctttaag	tcagaatatg	13680
taccttagag	tttttatatta	tgcttttgct	taccattaat	atgtctgcac	ctgctcttta	13740
gaagttaata	gagagtaaaag	tcgtctttat	gtctttcagt	gcttacttat	atttgggaag	13800
ttgagaaaaa	tttttaacat	cattattgat	atatatatata	atatatatata	atatatatata	13860
atatatatata	atatatatata	atagataatt	tttttttttt	tcttgagacg	gagtcctcact	13920
ctgtcgccca	ggccggagtg	tgggtggcgt	ctccactcaa	tgcaagctct	gcctcccagg	13980
ttcaagcgat	tctcttgccct	cagcctcccg	agtagctagg	atacaggctc	ccaccaccac	14040
gcctggctaa	tttttgtagt	tttagtagag	acagagtttc	accatattgg	ccagctgggt	14100
ctcaaaactcc	tgacctttgtg	atccgcccac	ctcgccctcc	caaagtgtctg	ggattacagg	14160
cgtgagccac	tgccgcccggc	tgaggtaaaa	tttaaagtgt	acaattcagt	cattttttagt	14220
atattttatac	tagttgtaca	gccatcacca	caatctaagt	ttagaacatt	ttcattaggg	14280
ggtgggagaa	attttactct	gctttttaga	ttaaagtctt	gtctggatct	aatcatttaa	14340
tcagacaatc	aggcagattg	tctgtgatta	gttttgccca	ttccagcttc	ttcattgggt	14400

gttaactttc	acaaataaaag	gctgctcaaa	gattagaaat	aacatttraat	ttgaatgtaa	14460
atgtgccata	gttttaaaga	tgggttttgg	gaatacagtc	aaatacatac	atttaaagct	14520
ctaattctga	agattatgta	aagaaaagga	aagaaatgta	gggagaggat	tgaaatgttc	14580
atggtataac	aatatctgaa	catccatctg	gtcacaccgt	tggatattga	atgttttgtc	14640
ctcctcaaat	tcatatgtcg	aaatcccaac	tcccaagggtg	atcgtattag	gaggtgtgggt	14700
ctttgggaag	tgattaggtc	atgaagggtga	agccttcacg	aatgggattc	gtgctcttat	14760
aaaagagaac	tgtgagaaat	aagtttctgt	cgtttgtag	ccaccagtt	taggatattt	14820
tgatatagca	gcctgcatgg	actgagacaa	ctatgagtta	ttatgatagc	ttctgttatt	14880
tcacctaaat	tcatagaagc	taatataatca	atatttatgc	tatgaaatat	ttcttaacca	14940
agctttgaat	atatttatat	ttttgtttat	ttttaaat	cagattccag	atgacctgag	15000
gaagagacta	aatatagaaa	tgcatgccgt	agtcaggata	actccagtgg	aagttacccc	15060
taaaattcca	agatctctaa	agttacaacc	tagagagaat	ttagtgaagt	caaatatata	15120
tgttacatca	aaattctttt	acacgttttg	taagatttct	agttgcttta	gctaagtaat	15180
aagaatgttg	tattcctttt	tgatacaaat	ctttttttat	tgtgttaaac	tatatataac	15240
ataaaatatg	ccatgttcgc	cattttttaag	tgtataattc	aaaggcatta	attacattca	15300
taatattgta	caaccatcac	cactatctat	atccagaact	tttccatcac	cccaaagaga	15360
aacttggtag	ccattaaaca	ataattcccc	gtccactcct	ttccccagtc	cttggtaatc	15420
tctaattgtat	attgtgtctc	tatgaattta	tatatcttag	atatttcata	cataagtaga	15480
agtatgcatt	tgtcttatgt	atctgactta	tttcatttaa	cataatgttt	tcaaggctca	15540
tctgtgttgt	atgtatcaga	atgttattcc	ttttcatggc	tgaatactat	tccattgact	15600
gcatatacca	catttgttta	tccattcatc	tgttgatgga	cacttggggt	gtttccacat	15660
ttttggctgc	tgtgaataat	gctacagtga	acattgggtg	acaagtatct	gtttgagttc	15720
ctcttttcag	ctcctttggg	atatacctag	gaattatgtt	taactttttg	agaagctgag	15780
aaatctttaa	taaatgataa	cacaaatact	tatatttgcc	aatgcaaata	tgaatatatt	15840
tggcttttaa	gagattgatc	attttgccac	gtggttgtaa	ttaaaaaaa	ttgtcccatg	15900
ttgtttcagt	attaatattg	tagcctaaaa	gagtgctaga	ctgttttact	ttttactcag	15960
ttaattcttt	ggatactggg	agagtcagga	aatgagatat	tgaacttaaa	gatctttgca	16020
ggtgggggtc	agtgggtcac	acctgtaatc	ctagcacttt	gggaagctga	ggtgggagga	16080
ttgcttgagg	ccaagagttt	gagaatagcc	tgggcaacat	agcaagacc	agactctaca	16140
aaaaaattaa	aaaaaaaatt	aagccaggcg	tggtagctca	cgcctgttat	cccaacactt	16200
cgggaggctg	agatgggtgg	atcacttgag	gtcaggagtt	ggagaccagc	ctggccaaca	16260
tggtgaaacc	ccatctctac	taaaaatacc	aaaattatcg	ggcggtgggtg	ctaactcctgt	16320
aatctcagct	actcaggagg	ctgaggcagg	agaaccactt	gaactgagga	ggtggaagtt	16380
gcagtgagcc	tagatctcac	cactgcactc	cagcctgggt	aacagagcga	gactctattt	16440
caaaaaaagt	aaaaataaaa	attagacaca	tgtgtgtgca	catgcctgta	gtcctagcta	16500
ctcaggaggc	tgactgaagt	gggaggatct	cttgagccca	ggagttccac	actgcagtga	16560
gctatgattg	tgccactgca	ctccagccta	ggcaatatct	caaaaaaaat	ttttttaaat	16620
agattattag	gccagacgtg	gtggctcatg	ccagtaatcc	cagcactttg	gaaggccaag	16680
gcaggcggat	cacctgaggc	caggagtttg	agaccagcct	ggccaacatg	gtgaaacccc	16740
atgtctacca	aaaatacaaa	aattagctgc	aatgtctata	atcccagcta	cttgggagcc	16800
tgaggcaagc	gaatcgcttg	aaccggggag	gcagaggttg	cagtgaagtgg	agactgcgcc	16860
actgcactcc	agcctgggcg	atacagcgag	attctgtctc	aaagaaaaag	gaatttgttt	16920
tcctgtcttt	atcgtagagg	gaggaaaggg	agaatggggt	tggaatgggt	attgagtgag	16980
ccacattatg	gtagatgtat	cactgggcat	agagaaaaag	agcatttaaa	acttttccgc	17040
ctaacagatg	tttcttcagg	ctacactgca	ctcattgtgc	taactgtaat	gtcaaatccc	17100
agacctgtgc	ctatagaaca	tgaacatcct	tcattggatt	tgtttgggtca	ggcttacact	17160
ttattaggaa	gatcagatgt	taaaataaag	gtgttaaagt	taagttcaga	tatgaggata	17220
attcattgac	attccttttt	ctggcagcct	aaagacataa	gtgaagaaga	cataaaaact	17280
gtattttatt	catggctaca	gcagtctact	accaccatgc	ttccttttgt	aatatcagag	17340
gaagaattta	ttaagctgga	aactaaagat	ggtgagtaca	tttgttattt	tgactttttt	17400
ttctatttaa	atagttgtac	atttttaatt	gttcttgcaa	cctgtcatac	ctgtgaacag	17460
tatgtgaata	gtgaaatata	attatgataa	ttaaacagta	gtttttatgt	attgaaaaat	17520
atctttggcc	gggtgcagtg	gctcatgcct	gtaactcccag	cactttggga	ggccgaggca	17580
ggcggtcac	ttgaggccag	gagttcgaga	gcagcctgcc	aacatggcgc	aaccctatct	17640
atacaaaaaa	atacaaaaat	tagcctgaca	tagtggtgta	tgctgtagt	cccagctact	17700
tgggaggctg	aggcagaagg	atcacttgag	cccaggaggt	ctgtgttcct	gccactgcac	17760
tccagcctgg	gcagcagagt	gagaccctgt	tggggggaag	aaaaaaaaag	tctttaactt	17820
aaataaattt	gacattttaa	atcttaaat	atttcatctc	tgtttcagta	ctaactctgc	17880

atttattact	ttcttttttaa	taggactgaa	ggaattttct	ctgagtatag	ttcattcttg	17940
ggaaaaagaa	aaagataaaa	atatttttct	gttgagtooc	aatttgctgc	agaagactac	18000
aatacaagta	atagcatgtt	attgaatat	taataaaaata	ctatttggtta	catatgattg	18060
ataataaagt	atgaagttcc	ttgtaacacc	ttgcattgtg	aagtgtatta	aaaacctgct	18120
aagagtaagg	aataacttga	tttaaaatat	tttattctgt	aatctcttta	aattatctgt	18180
acaaattatt	gacttaacct	aaatttaaaa	atgaatgcct	tagcacaatt	aagttccaag	18240
aatagagttg	atcatgttaa	ctggtaaatg	gatcatgatt	taaaattctt	ctaggattga	18300
aacaaatgaa	aacgtagttt	taagggtttg	atttttttaa	ttcctatttt	tacatgcaat	18360
tttactgcac	aacctatctt	attttgacag	ttcttaaatt	cgcaactctt	cagaaatatt	18420
atcagatcac	ttttctttgc	ttocataagt	ttttttatta	ttatattatt	attttttttt	18480
tttaaaagac	ggtgtctcac	tttgtcgcct	aggctggagt	gcagtggcat	gatcatggct	18540
cactgcagcc	tcgacctccc	aggctcaggt	gattctccca	cctcagcctc	ccaagtagct	18600
gggaccacag	gcgaatgcca	tgatgctgg	ctaatttttg	tatgttttgt	agagatagg	18660
tttcaccatg	ttgcccagaa	ttgtcttgaa	ctcctgggtt	caagcagttg	ttctgccttg	18720
cccacccaaa	gttgtgggat	tacaagtgtg	agccactgog	cccagctatt	ctagaagtat	18780
tttaagagtc	atcttttttt	tttttttgag	atggagtctc	actctgtcac	ccaggctgga	18840
gtgcagtggc	acactctcgg	ctcactgcaa	cctccacctc	ctgggttcaa	gtgattctcc	18900
tgccctcagct	tccttagtag	ctaggattac	aggcgcattg	caccatgccc	tgctattttt	18960
tgtagtttta	gtagagacga	gatttcacca	tgttgccag	gctgctcttg	aaactctgac	19020
ctcaagtgat	ctgcccctct	cagcctccca	aagtgcctgg	attctaagtg	taaaccacca	19080
caccagcca	agagtgggtct	ttttacaata	ttattttttg	attaggacat	tcattcttgt	19140
cataaaattg	aagatactct	agtcatttag	aatttcattg	ttttggaaat	agacattgtt	19200
tctttatttt	tgaaatgtta	ttgaaggaat	accatttgga	gaagatacaa	atgtaagaat	19260
tgtgaaaagg	ataattgtga	cacaaatcaa	aattatagat	aaaaatatac	ctgtaaaatg	19320
tattaaggca	ataacattct	ttctgcttgt	tgaccataaa	tatttatatt	ccctggatgg	19380
gtacattggt	attgtcaagg	gtgtttaaat	aatgatcttg	catgcataat	ttattctctc	19440
tggtataaca	gaatcagcaa	tttagttttc	tgggaccoga	gaaaaacatg	caaaagacat	19500
actttgaaat	gtaaaactga	tttttccctg	caactgtagg	tccttctaga	tcctatggta	19560
aaagaagaaa	acagtgagga	aattgacttt	attcttccct	ttttaagct	gagctctttg	19620
gggtaagaag	ttatggccaa	actagcatgt	tagacatgtt	tttaacacta	tatctggcag	19680
agttttcaat	gtaaatatta	aagttagatgt	taattgtcaat	aagtgatctt	aataatgcat	19740
cagtagatat	tttttcaagg	attgtctcta	tcttcacgcc	tagcttataa	tttgccctgt	19800
cgtctttttt	tttttctctt	tatttttatg	tttttatcca	tccctgggtg	taggggataa	19860
ccttgtcttc	ttcgataaca	agaagtctga	agcttattag	aaattttact	ttgagaattg	19920
atcgatgaga	agaaagcaac	tagatatcac	gtggatcata	tatgcttgaa	taaaacaata	19980
attcttagaa	caataaata	catttttaaaa	gttaaaagcca	aaaacattag	ttgaaatgtt	20040
aaaaatattt	caaattaagt	tattccttca	ctgtcttgta	ttactgtaat	aatttggatt	20100
atttgtgttt	ttctcaactt	ttaaaacaaa	tattttaaaa	attcctcttt	tgattaagta	20160
gggctagata	aatataaaa	aatatttttt	aaactcctct	taatttccat	atttcttata	20220
taatatgaga	atctcttata	aacactacct	cttagaagtc	tcacagaaag	ctttggtaga	20280
tgtagtagta	gggatttgat	ttcttagaat	ggtataatct	gtaaatgttt	tagtaaaagg	20340
attaaacgat	aaagtcaaaa	tgtttatagc	acagtgttta	ttaatataaa	ataaaatctc	20400
tttttttttt	tttgagatgg	actctcactt	tgtcactcag	gctggagtgc	agtgttgcaa	20460
tctcagctca	ttgcaacctc	cgccctcctg	gttcaagcaa	tccttccgca	tcagcctcct	20520
aagtagctgg	gattacaagc	atgcaccacc	acacctgcct	aattttttgt	atttttagta	20580
gagatgggg	ttcaccatgt	tggccagget	ggtctcaagt	gatccgcctg	cctcagcctc	20640
ccaaagtgct	gggattacag	gcgtgaacca	ctgtgcccag	cataaagtaa	aatctcttca	20700
gactctcatg	tgatcatgta	aagtggcagg	cagtcacagt	caagaagtag	tttaaagttc	20760
atgtttgtaa	aatataatct	acagattgat	actggatttc	ataggtaatg	tttaagagaa	20820
aataagtttt	tagttatcct	cagtacttca	aaagcaccga	tttatgatta	tgttgattac	20880
taaactaaat	catttggggg	ctagaggtgt	ttttttatgt	gttaagattc	cttaaggagt	20940
tctattagg	caaaactttt	agtaactgca	tatttttaaa	gtaataaaac	taatttttaa	21000
agcttggagg	ctgggcgcgg	tggctcacac	ctgtaattcc	agcacttttg	ggggccaagg	21060
cgggtggatc	acttgaggtc	aggagtttga	gacgagcctg	agcaacatgg	tgaacacttg	21120
tctctactaa	aaatacagaa	attagccagg	tgtggtgggtg	ggcacctgta	atcccagcta	21180
ctcgggaggc	taaggcagga	gaattgctcg	aacttgggag	gcagaggttg	cagtgagccg	21240
agatcatgcc	actgcactcc	agcctgggtg	acagagcaag	actccgtctc	aaaaaaaaaa	21300
aaaaaaaaaa	gcttgaagtc	agattcgaca	ttaatcagta	tactttctct	caagtagggg	21360

acaattttcta	agatttttagt	cttttaaaat	ttattaacta	gtctgagcat	gggtggcttgc	21420
gtctataatc	ccagcacttt	gtggggccga	ggcagatgga	tcacttgagc	ccaggagttg	21480
gagactagcc	tgggcaacat	ggcaaaaccc	cgtctctaca	acaaatgcac	acacaaaaaa	21540
cccaatcagc	tgggtgtggt	gttacactcc	tgaagtccca	gctactcggg	aggctgaggg	21600
aggaggatca	cctttgccag	ggcgtttgag	gctgcaggga	gctgggttca	caccactgcg	21660
ctccagcctg	gatgacacag	caagcccctt	tctcaaaaaa	aaaaagataa	aaaattaaat	21720
taaatttaatt	aactacactg	ggaaggcaaa	attcagcatt	tttttatagc	taaattttat	21780
cctgcttcag	tcttttatca	tgtaaactatg	tatatTTTT	acagaggagt	gaattcctta	21840
ggcgtatcct	ccttggagca	catcactcac	agcctcctgg	gacgcccttt	gtctcggcag	21900
ctgatgtctc	ttgttgcagg	acttaggaat	ggagctcttt	tactcacagg	aggaaaggta	21960
agtgggttaag	gtgtgttcat	ttttctgtaa	catttaataa	cttttcattt	atcttctctt	22020
gggtttttgac	catctattat	atagggtggg	ttttgaccat	ctattatata	gggttttatac	22080
gacatatgga	aagcattcat	ttattcacta	atatttctgt	gtgtctgctt	ttaggtgttg	22140
ggggagtgat	gacgaataag	actgatgttc	tccatgccct	ttttctgtgt	cagttgatac	22200
aatttatatgg	tttttctttt	ttaggctatt	agggtgtgat	agggttgagt	aacttacaaa	22260
tgttgaacca	gccttgcata	cctgtgataa	ataccacgta	gttgtgggtg	atcattcctt	22320
ctacattgct	gagttttatc	tgctaattgtt	ctgttgagct	tttgtccatt	taagtttgaa	22380
agtgttttag	ttgcagtttt	ctgtttttgt	gtgtgtctttg	tctgggtttg	ctatccgtgt	22440
aaatctggcc	tcataaaatg	agatgggaag	tattctctcc	tcttcttttg	tttttttggg	22500
agaggttgta	taaaattgag	gctgaatctt	gggtggtgcc	acaatgacag	gaactatttc	22560
tgtgactgaa	tatatggga	attcctataa	agcaattatt	ttctagggaa	gtggaaaatc	22620
aacttttagcc	aaagcaatct	gtaaagaagc	atttgacaaa	ctggatgccc	atgtggagag	22680
agttgactgt	aaagccttac	gaggtatgag	tatggtaaca	ctctatataa	atcccttttt	22740
cattagaaaag	acaggaatgt	tatacataat	gctgtcaatc	taataaatac	acatatcatc	22800
tagtcttttaa	cttttctgtt	tatcatttag	tcattaaaaat	ttctttggct	ttctaattgt	22860
tttgataaaa	tttctaaaac	tctccatatt	taatggaggc	ctattttttt	ttctagccag	22920
aactttttgt	agactacatt	tctggaagtg	ctcactgaca	ccactctgaa	aaatttagtac	22980
ttagaatata	ctctaattgg	tataaatgat	ctctgaattg	ctatggaaaa	ctgggagaat	23040
ggttgcttca	ggggagagaa	agtaggaggc	tgtggacagc	aatgaggaga	attacagttc	23100
accatataac	acttttgtac	ttttaaaagtc	cttaacattt	acattattat	ctattcaatt	23160
aaaaaatatt	gggaagattt	tactttgaac	agttaatttt	tcccccatgg	gtaccgctgt	23220
catatagtcc	caactaatca	tgaacttgtg	tatttccctgt	tctttgtaaa	tttaaacttt	23280
gtaactcacc	aggaagtttg	aagccaaatt	tgtgtttcaa	atatagcaac	tccaggatct	23340
ctaggcgat	gcatttgcatt	ttgattttta	atgaatcttg	atcccttact	ctcacttatg	23400
ttttccca	tcctactttt	tttattttgt	tgtaaagccat	ctaaaattct	caatgggatg	23460
aaactgggt	taaatgaata	catgcataca	ggaattatag	tagcatattc	cttttctttt	23520
ttcttttttt	ttttttttga	gacagagtct	tgctctgtag	cccaggctgg	agtgcagtgg	23580
tgcatctcg	gctcactata	gcctccacct	cccagggtca	agcaattctc	gtgcctcaac	23640
ctcccagta	attgggacta	cagggtgcag	ccaccacacc	tggctaattt	ttgtattttt	23700
tagtagagat	gggttttcac	catgttggcc	aggctgatct	caaactcctg	acctcaaagt	23760
gatctgcctg	ccttggtttc	ccaaagtgc	gggattacta	gcataagcca	ctgcacctgg	23820
cctccttttc	tgagttttat	aaaatttgat	actttactgc	acgctttgag	actgtattaa	23880
ttgaaccatg	ttgatgaaca	agtttttgtg	atgggtatat	taataaaaata	tagatcaa	23940
ttttatagtt	aaatcaatat	cgagcttttc	tagtgctttc	aaaaggacaa	cctgaatttt	24000
cccagcactg	aaatgatact	gaaaccattt	catatcttct	gtattaagga	aaaaggcttg	24060
aaaacataca	aaaaacccta	gaggtggctt	tctcagaggc	agtgtggatg	cagccatctg	24120
ttgtcctgta	ggatgacctt	gacctcattg	ctggactgcc	tgctgtcccg	gaacatgagc	24180
acagtcctga	tgcggtgcag	agccagcggc	ttgctcatgg	taaatgcac	caccactggc	24240
ttaaggtctt	gttcttttgt	cagtcagcat	ttttagtctt	aacaataaat	ctactctctt	24300
cagagaataa	tatatgtgtt	atgttaagtg	ttgtgtttga	ggccccctgat	ggcattctac	24360
agttgtccta	tagactgtaa	tagcaaaatt	ggtagagtaa	aaacagtgtg	aaaattctgc	24420
aacttcatgg	ttagtccctt	agggtttttc	attctccctt	acttattgtt	taattttacag	24480
atttactctt	ttgttcattt	gacaaaatgc	tgtcaaatgc	ttgtgcacag	tctgtattct	24540
caaattctag	gagaaaaaga	agggtgaaca	gtattagcgc	agaacgatac	taataatgat	24600
ggctactgtg	tatgagtagc	cagccctttc	ttggctttct	tggattgctt	tgtattctac	24660
atgaagatat	tccctgggct	ttacaggtca	ataaatggaa	attcagagag	attaatttga	24720
ccagggtgac	caacaaggag	atgacagcat	acactatgcg	agaagtatac	acagagtagt	24780
gtaggagcat	ataacctaaa	ctgggggtga	ggtgggataa	ggagttatca	gggaaggctt	24840

tttggaggag	ttgacaactg	agccgagttt	tgatggaaga	gtagaaatta	gcatgaacca	24900
atttcatgct	aataaagaag	caaaggaagc	gtggtctaca	ggcaaaagca	cagagggtaca	24960
ggaagtaatg	atatgttggg	gaataccctg	ttgactggag	cttagagtgc	aaggagagga	25020
gtgctagggg	ggtgagggtt	gaggggtttg	cagcattgac	ttgcttcaag	gttcttaaga	25080
gctgaaatag	atataaaatg	caactaagag	tggcttggat	tattattacc	tagtgtgtta	25140
atctcaaatt	ttgaaatcta	tagcatctat	aggactgggt	ttactaatct	tacactcgat	25200
ctgttactgt	tcttataacta	gatctattag	tccagtgttt	aagggagtgg	tgcagatttc	25260
taggtcagga	caggactcag	atgtacatta	ttaatgccta	tttcagttct	gaccttctca	25320
tatgaaacct	tataagacct	ggggtaggaa	gagattgttc	tggaaagtc	aggaatatga	25380
actgtatfff	gtttaacaaa	caatacagta	tggaaattta	tcaccttcc	agaatatffa	25440
tttcagagac	aaatffffat	cattcgttca	tttatttcat	aagatccacg	agtagggaa	25500
ctcactagac	attgctctga	gtatatgggt	tgagtttgca	gtacctcttg	tgtctccatt	25560
agatttatta	ggctctcaat	agataaatca	gggaataact	agatggatto	attttttaa	25620
gacatgaaag	agcgatacca	tacatactgc	accttaagg	tcaaccttag	agtatcatta	25680
tttttaata	atgtataatt	tttaaatttc	atgtttactt	ttcctaagct	tttgactat	25740
attgctta	tccagctttg	aatgatatga	taaaagagtt	tatctccatg	ggaagtttgg	25800
ttgcactgat	tgccacaagt	cagtctcagc	aatctctaca	tcctttactt	gtttctgctc	25860
aagggtttca	catatttcag	tgcgtccaac	acattcagcc	tcctaatacag	gttaatacact	25920
acttgtaagg	attattgaat	tatgtccctt	ttatagaaat	tatttttcaa	ttttattagt	25980
aattcgtggc	tttaaattta	tgttctctt	aatgatttta	aggatatgta	agtcaacatt	26040
tgggtgcatat	tgtgctagag	gcataaatta	taatttatag	ccacctgaaa	tgttagtatg	26100
cgctttccaa	gaaaatgact	tttttgaaaa	tggatattct	ttgaatgaga	aagaacagag	26160
agaaatagat	agatggcttt	taaacacttc	attaattaaa	cttttttttt	ccaccatcac	26220
ataatggcac	ttagtccctt	ttgggaactc	atgaggggtt	tagtggtagt	gagctgaaag	26280
aaatatgttc	caggactggc	aaacatattc	taaattcttt	aaaattttca	cctagcatct	26340
accctaaata	ttcagaccct	gtgctagtta	actgctattg	aagaacaaag	gtattatata	26400
tattattaag	gataatagaa	tggatattga	gatattgggt	attgaatatg	aatatgtttt	26460
gagaaataag	ttttatagga	acaaaaaaa	aattctttaa	ggaaccatat	attactaaaa	26520
atgcttctta	ttggagaaa	aaatgacaat	catattatta	tgtgattttt	tcacaacttt	26580
attaagatat	aatttaagta	caacaaactc	acataaagtg	tacaatttga	tcagttttta	26640
catatgtaga	tgccatgaaa	ccatcaccac	aattaaggaa	acaaacattt	tcactactcc	26700
agaagtctcc	tagccctttt	actacccatt	cctcccctgc	tccatcccca	gacaactacc	26760
aatttgcttt	ctgtcactat	agatttgtca	acctgatttt	ctccaaatat	acattcaaaa	26820
atatacagtt	gaatacaatt	ggaaattcga	attttgtgtt	tttttcttta	ggaacaaaga	26880
tgtgaaattc	tgtgtaattg	aataaaaaat	aaattggact	gtgatataaa	caagttcacc	26940
gatcttgacc	tgcagcatgt	agctaaagaa	actggcggtg	ttgtggctag	agattttaca	27000
gtacttgtgg	atcgagccat	acattctcga	ctctctcgtc	agagtatatc	caccagagaa	27060
agtatgtttt	actattaaaa	cctgaacttg	gaatcttctt	tctatttgtg	agaaatgtaa	27120
ttgtagtaag	acaagaatta	aatatattcc	attgtagtat	ttgaataagc	agttatttga	27180
gtagaaaatt	agtgtttcca	gctaagatga	tggcatattt	tgaatttca	tatagtgaat	27240
ataactagta	aaagaagttt	tgtttatttt	taaacagaat	tagttttaac	aacattggac	27300
ttccaaaagg	ctctccgcgg	atttcttctt	gcgtctttgc	gaagtgtcaa	cctgcataaa	27360
cctagagacc	tgggttggga	caagattggg	gggttacatg	aagttaggca	gatactcatg	27420
gatactatcc	agttacctgc	caaggatgtg	ttaaaaaaag	aaaaagtga	tacttactcc	27480
cagaagaacc	actgtattat	tggctttggc	tttatgtgtc	agcttgccca	atctccgtgt	27540
gagtcaacaa	gtgtttactg	agttaccaaa	taaatgtctt	aacactattt	taggtacttt	27600
aacaaaattt	aattttatta	attaattttt	tattagaatt	gagacctcac	tctgtcatct	27660
aggctggagt	acactcacag	ctcactgcaa	cctcaaaactc	ctggggtcaa	gcaatcctcc	27720
tgcctcagcc	tccccagtag	ctagaactac	aggcatgaac	caccatgccc	ggccaaactct	27780
ttaattttct	tagagacgga	gtcttgctat	gttgcccagg	cagacagatt	ttaatgtgta	27840
tgatgcagtc	tttgatgata	agaaacttat	aatggaaaagc	tgaggtgata	gttacagtaa	27900
atacatftttg	atgtataatt	ctgtttgctt	taatcattca	aattgtagta	aagcaagatg	27960
aactgtctgc	tgggatttga	gcagaaatgg	ataggaataa	actaggaggt	agaagagtta	28020
tcaaggttca	caggactgat	gggtgaaagt	agatttccag	acccgggatg	tcagtccttg	28080
aaaagcagac	ttggcaggca	tagacgaggc	agatagcagg	ataaaggaga	caaatgtaga	28140
ttgttcttca	gaagatcaga	tggtagagtc	taggaggtag	tgtgttttaa	tcagagatct	28200
gagaggcaaa	gatcatttga	tgagatcagg	gacctatgca	aaggagttag	aaaaaaaaact	28260
gggttaagga	gcctgctgca	tggcaactcc	tgggaacagt	ggccactggg	gcctgggaca	28320

tggttgattgc	agcccaggac	tggttaaaaco	agtgtgagag	aacatgggta	tggaagtaact	28380
agctagcagg	atcatgaccc	cgatgctggg	atggggcctc	aagcattagt	acatggagat	28440
tcagtacatc	cagatgcagt	acatggagac	tatatgcgta	actgctgact	ttgggcttct	28500
ttcagattgg	agcagaggta	gaggtgagtg	ggaatattct	caatagaggg	aactaaatag	28560
gcatacctaa	taaaggagac	caggatattg	cagacagtag	cctcatgttt	ggctcacctg	28620
ttcaaaaagt	tctcttggtc	ttgagcagtg	gtgccttaaa	aggtaacttg	agaagcagtc	28680
gattattttgt	tcagcctgga	gactcttggt	atattttact	atctttgatt	gaatagattt	28740
aaatgtacac	agctctcata	acttgcccca	tgaagcatat	ccatgaaagg	cactatactt	28800
gttaaaagat	tggtttgtac	tttttaaatg	tagtactttt	aataaaacag	gaaaaataga	28860
agttctgatg	cagtttatatg	cattttatat	agaatgtgtt	cttaattgga	aaaaatttgt	28920
cgtagttcct	ttgagttcat	ttacagtttt	tagtaggaat	tgtattttct	actgtttgtac	28980
ttgctgtttac	taaagaaaaga	tggtcgtgat	taccatctga	attttttttc	tatacatatga	29040
tcttttagctg	ctacttagtc	atttctgttt	agacttgagc	tctttttcat	attttttttt	29100
tttgttttctc	agtatccaga	attattttgca	aacttgccca	tacgacaaag	aacaggaata	29160
ctgttgatg	gtccgcctgg	aacaggaaaa	accttactag	ctggggtaat	tgacagagag	29220
agtagaatga	attttataag	tgtcaaggta	tggtgtctac	ttatcttctt	tttttatctt	29280
ggtaaaatta	acataaatgc	agtttagccat	ttcaaatgt	aaattcactg	gcattttagt	29340
cattcacaaat	gctatgcaac	caccacctct	ctctaatttc	aaaacttttt	cattccactg	29400
ctcctcttgc	ttatcccttg	gcaaccattc	atctgctttt	tgtctctatg	gattttgcctt	29460
ttctgtatat	ttcatataaa	acaaatcatg	caatatgtga	ccttttttgt	ctggcttctt	29520
tcacttatgt	aatgttttca	tggttcattcc	aggtagtagc	atgtatcagt	acttcatctc	29580
tttgcatgac	tgaataatgt	taccataact	tgtttatcca	cttatcagt	gtgaacattt	29640
gaattgtttc	taccttttga	ctattatgaa	taatgttgct	gtaaatatcc	atgcacaaat	29700
ttctccacgg	atatgttttc	atttctcttg	ggataaaact	gaggagtaga	attcttggtg	29760
cttagggtaa	ttctctaact	tttcaaagaa	ccaccaaact	gtctttcaca	ccaactgcac	29820
cattcccact	agcagtgtgg	gggtttcctg	attctccaca	tctttaccac	caccattatg	29880
tttctcaatt	gtgggctagt	ctcacatttg	gaaagctagt	gggagcagcg	atccatctat	29940
taaaagttgt	atgaaattga	gtaatgagcc	acctctctct	tgtagggtct	attatgttct	30000
tgcttaaggc	aatcttcatg	cattgtgaac	agaattatac	ataaatgctc	agataaaaagg	30060
gcaaaccatt	cttaaaggga	gtagacaact	agaggcagga	gaccatactg	aggcaggaag	30120
ctgggggtttt	tatgggtctg	ttacttttga	ctatatctca	ccattgcttt	tgtcaaagtg	30180
agactaggtc	taagtttttt	tcagggtataa	ggtagtggtg	gtaattaaagg	ggcatgctag	30240
cagatcattt	tgggtaaatg	ttcacagtc	accactgggt	tgtcattgtg	gtcgcagatc	30300
cagtagctta	gctgtgtaat	ttcagacatc	agcaatatta	gtttaacaaa	gggcaattag	30360
attccaagac	aaaggaatcg	tgtattattc	tagccttatt	caaacttgat	ttataaatca	30420
gttttagtaat	ttattttatt	gtttctgtat	ttatttttat	ttctttgaga	tggagtctca	30480
ctctattggc	caggctggag	tgtagtgtat	caatcttggc	ttactgcaac	ctctgcctcc	30540
tgggttcaag	ctattctcct	gcctcagcct	cccgagttag	tgggattaca	ggctaatttt	30600
tgtattttta	gtagagatgg	ggtttcacca	tggttgccag	gctggtcttg	aactcctgac	30660
ctcagagtga	ctgcccgcct	tggttcacca	aagttctggg	attacagacg	tgagctaccg	30720
tgcccagctc	agtttagtaa	tgtataactg	ggttttaccc	agttgtaaat	tactcttttg	30780
tcgtgttttt	ttgagaactg	gcaatgacgg	agaaactaaa	agtgccaggc	tgttgccctg	30840
ttcctgttat	tttgcccttag	tttttttttt	tttttttttt	ttctctgaga	ctgagctctg	30900
ttgtgttacc	aggctagagt	ggagtggcat	gatctcggct	cactgcaacc	tctgcctcct	30960
gggttcaagt	gattcctgcc	tcagcctccc	gagtagctgg	gattacaggc	gcctgccacc	31020
gcacccgggtg	aattttttgta	tttttagtag	agacgggatt	ttaccatgtt	ggccagggtg	31080
gcctcgacct	cctgacctca	tgatccacca	gcttcggcct	cccaaagtgc	tgggattaca	31140
ggcgagaacc	accgtgcccg	gtcttgccct	agttattttt	tgttccctcc	tctagtctca	31200
tagttctctg	actgtattga	ggaaatgtaa	ttaaatatta	ttatgttaat	agatatttat	31260
gtgggttgaat	attagaaatt	ccttattttt	gtcacatatc	ctgatcagta	gttggtcttc	31320
tggagatagt	gattttttcac	tagagatgac	tttaggacct	attcagggtt	tttttaagat	31380
cccaatttaa	ggaaagacta	ttctcattat	tgatttttgt	atatgcaggg	aaattttatt	31440
cgaaagggtt	ttcagttggc	ttttagggaa	gattatatat	tctctttttt	tttttttggc	31500
cttttcccac	atgtttctaaa	aatgatatat	tcttttaactc	ctatgaaaaa	acattgtttc	31560
agtaattgaa	gatgctgatt	aaagtcatat	ctctacacat	tttttaaaat	ttgagataga	31620
tgggactttg	tcccttctta	caccattcac	ttattcactt	ggaaaaacta	ttatccaata	31680
cttatgtggc	agacactgtt	tctggcacia	gggattcagc	agtgaacaaa	actgcctttt	31740
tggagtttac	attctactag	tggaaagcga	caacaagcag	atagacacat	tcagtatata	31800

attcactgtc	agatgggtgt	ggtaagtcct	atgtaggaag	aaaagcaggg	taaggaggct	31860
tggagtaact	ggagtgtgtc	atagatggac	ttgtcaggaa	aggggtttctg	aagagggtgt	31920
atttgggcag	agatctaaat	aaaatgaagc	aacaagccat	gagaatatcc	gggggaaaaat	31980
gttctgggca	gaagcatcaa	gcatagaact	tgtggtatga	tattttattct	agcacacatt	32040
aattttaaaa	atgtataaaa	gacatccatt	taatcatatt	aaagattttcc	atgattcatt	32100
tagacttagt	cagaaaaccaa	atttatatatt	tctttttttaa	taatttttatc	tcaactcttta	32160
ttttacccaa	tagggggccag	agttactcag	caaatacatt	ggagcaagtgt	aacaagctgt	32220
tcgggatatt	tttattaggt	tggtagccta	tgaatgtttt	taaagtaact	gactctgtta	32280
ttattttatca	atcagtgcct	tttttggctc	tgttttttga	agaactgata	tttgaaacct	32340
gtggtttatg	tgaattatta	ataagctaga	ggacgtggat	tctctatttc	atcaaataat	32400
acaaaacatt	ttagatatta	aatttttgaa	attatttggg	tttggttttac	aatagaaata	32460
ctcctcaaa	tggaatcgaa	gtggttattc	aaagaaaatct	cagagtagat	tcttatatga	32520
agcaaaataat	tgccccta	ttatctctaa	attttgttaag	ttctaaattc	ttttttcccc	32580
cagtttctaa	tttatctctt	ataagtcaag	agtccatctg	gccaaatttaa	tttcagtgtg	32640
tgtaaactatt	ttgcatatat	taaaaaactg	tatatgaata	cagaagatgg	tattttaagga	32700
tgaaaaataat	tattcaaatg	tgatagcatt	atggggagtt	ttaaaataaaa	agttactgtt	32760
ttatttcttcc	aaaaatttta	ttataaagta	tacagttaa	agaatataca	taaaatacat	32820
atgcagctta	aggaagaata	ataaaatgaa	tacttcattgt	attcaccacc	gagtttacca	32880
ggaaaaagca	taaacaaaat	aaacctcttc	cacgtaattc	ctgggtttaa	gagaagttat	32940
agtggaaaat	atttgggagc	aaacgataat	gaaaatacta	tccattaaaa	ttgttagatg	33000
ttgcaaaaact	gatttcaagg	aaaatttata	gtgttaaatg	tttagaaaaag	aaaaaagggt	33060
agaagttaac	cacttatgta	tctatctcat	gaaattagga	aaattataga	tataaactaa	33120
aaaatatgtt	aaaagggaaa	taataaagat	aagaatgaag	tttaatgaaa	cacaaaaacag	33180
agaagctcac	aaagccaaga	tttatttttt	gaacaccgag	tacaattgac	aaatctctaa	33240
caagtttgat	taagaaaaaa	gaaagcatga	ataaacaatt	ttagggataa	aaagggaac	33300
atcgctaaag	atatcccaga	aatgtaaaag	ataataagg	aatattatga	aaatattcat	33360
gccaatacat	ttgaaaactt	aggtgacata	gacaaaaaca	aaattgacca	aaattgagca	33420
aaaaagaaac	aaaatctgag	tagtcctgta	acttagtaaa	aattgagtta	gaaaagttaa	33480
agaagtcttt	acacaaatca	aacatcagac	tcagttttct	aggagagttt	tgccaaacat	33540
tcaagtagca	gataattctg	gtctattttt	ggcccagaa	gatataattt	acttgccatg	33600
catttaaatga	gatagctgtt	gatttttttt	aatcaccgtg	acaggtgttt	tatattaggt	33660
gttattcgcc	agacatctag	tccacctgtt	gccagatatg	gaattaatat	tcacttattt	33720
tgaattaaaa	tttggttaata	aattaataaa	acaaagtcaa	agttcaaatt	attaaaaaag	33780
taaaagaaaat	aaaatatatt	ttatagagag	cccttacaaa	acagtaccaa	cataatgagc	33840
tttccaaatt	tgaatgggc	aaaataaatg	ataaggcatt	tcacaaaaga	aggaagggtg	33900
gccaaagaat	atataattat	ataaaaaatg	ttacttgtta	taggaatcaa	aagtgtttga	33960
cttattgact	aagagtcagt	ttttgttttg	atccctgtta	gtctatccag	aaggcatggg	34020
tcttaataaaa	caccttgacc	tcaacagttt	actgaataca	agggttaatt	catatgcctt	34080
gccttcttta	agggtttgtt	gtaaaagatta	aaataaatac	ataaataatat	ataaatacat	34140
ttatatgtat	ttatatgtaa	ttacatacaa	cttgcccttct	ttaagggttt	gttgtaaaaa	34200
ttaaaagaag	tataataata	tataataata	cataaaaata	atacattcat	atatgtatat	34260
gaaatcactt	tgccaactat	gaagcctgat	tcaaataatga	aatgttggtt	gtttttccca	34320
gagcacaggc	tgcaaaagccc	tgcattcttt	tctttgatga	atttgaatcc	attgtctctc	34380
ggcgggggtca	tgataataca	ggagttacag	accgagtagt	taaccagttg	ctgactcagt	34440
tggatggagt	agaaggctta	cagggttaata	attataaata	cagaaataga	atgttataac	34500
aaaatgtcat	catgtcatca	gatttttggt	aaaaaatgtt	cttttttctc	ctagggtgtt	34560
atgtattggc	tgctactagt	cgccctgact	tgattgacct	tgccctgtct	aggcctggtc	34620
gactagataa	atgtgtatc	tgtcctctc	ctgatcaggt	gacaatttca	tatttagagt	34680
ccaaaaccca	acaaatgcta	cactctttcc	ttgtgagctt	tacttctgcc	aggtaatggc	34740
aattgtcctt	agaagaccag	ctttcttagg	gaaaagcttt	agccactgtt	tgctcaaagc	34800
ataaaaagat	tctgaattag	atgcaaaagc	tttttttggc	ccagtgcagg	tctgaaaact	34860
ttgtaatcct	tctgtgttgg	ctgattgggg	aaaaaaaaat	gcaagaaacc	taatgtatta	34920
tattttcaca	ttatcttctg	ttcaaaagatt	atacacttcc	attatcctgt	caaaaaaaa	34980
actctgatac	agaatcaagc	atgtgaatcg	taagcatgta	agcagggtttc	atagagataa	35040
tttttcaact	cttccttgtc	ctgtgttggt	ccaactctta	ttctccaatt	tagaagcaaa	35100
caaataaatg	aatgaaagaa	cagatagaca	aatgaatagt	caaagggtata	aagtatctgt	35160
atatatgtta	catgtagcta	ttattttaa	tatttagatt	ttccttttga	aataccttct	35220
tggcacactt	gcctaaatct	agaaaataag	cactgtgtga	ataagaaatt	atttacactg	35280

aatatttttgt	aggttttttgg	gttttttgttt	ttcagacaag	gtctcacttt	gtcaccagg	35340
ctggagtaca	ctggtacgat	cacaactcac	tgcagcctct	atggcccagg	ctcaagcaat	35400
ctccccacct	cagcctcccg	agtagctggg	accacaggca	cacgctacca	tgcccagata	35460
atttttattat	taatttttgt	atagagatgg	ggtctccctg	tggtgcccag	gctttcttga	35520
actccagggc	tcaagtgatc	ctcccacctc	aacctcccaa	agtgttgga	ttacaggcgt	35580
gagccaccat	gcccagcctt	aagagtgttt	gatttttcatt	catttttccta	tatatattat	35640
ttctgttggg	gaaaaaatc	caaggaagat	aaatagtagg	ctgttggtac	atttctcaac	35700
ttacttataa	agcttttttag	atatataagg	ttaatttatg	aagaaaatca	taagatacac	35760
aatttaagat	aatattttta	atttttat	ttatttgtta	aataaatttt	tctcctttca	35820
ggtgtcacgt	cttgaaattt	taaatgtcct	cagtgaactct	ctacctctgg	cagatgatgt	35880
tgaccttcag	catgtagcat	cagtaactga	ctcctttact	ggagctgatc	tgaaagcttt	35940
actttacaat	gcccatttgg	aggccttaca	tggaatgctg	ctctcgagt	gactccaggc	36000
aagttatatg	aggaagtgtg	tatgacattt	tatgagtgat	aaaagaagta	caatgtcaaa	36060
atttccacct	taaaaaatgc	tatttttttaa	acaactttgg	taaaactgta	tagaaacata	36120
aattttacctt	tagttgaatg	ttccatagtt	ggaatatggg	ttttgcagag	aattttataat	36180
tatgaagttt	gatgtctgtt	tcttttaacat	taccttaata	ttggcaaaaa	catgttggtg	36240
tttgcaagga	tattattttaa	attgggatac	catgaatttaa	atactacaaa	caaaaataat	36300
tagagttttt	tgtttgtttg	tacttttaact	tttaaaaaat	aatcagttaa	agttgttgtt	36360
ttgaagctca	cattgttcca	atctggccaa	taggagcccc	ttttgtatgg	ctcctgtatc	36420
tttatgacat	gtcctcatca	ttcttgaatc	acttctcac	ttccagatac	agtaagttat	36480
tcttgccag	gtgcagtgg	tcacgcctgt	aatcccagca	ctttggcagg	ccaaggcagg	36540
aggatcattt	gggcctagtt	tgagaccaaa	tcattggttg	acaaactgta	cccactatgg	36600
acaacagagt	gggatcctgt	ctctgtgaaa	aattttaaaa	ttagctgggc	atgttgccac	36660
atacctgtag	tcctagcttc	ttgggagagg	ctgtggcagg	aggatcgctt	gagtaaattcc	36720
aggatgcagt	gagccatgct	tgtgccactg	cactccagca	tggtatgacag	aatgagacct	36780
tgcccccaaa	aaagaaaaat	attccttggtt	tatcctgtac	tttctgtatc	ccagccctag	36840
catcagcctt	ttctctaaaag	acagtattat	gatttttaata	tttacagtag	atattttgaac	36900
tgttacatta	tagactttac	catatatattt	ctaggaagga	ttattctatt	actcctcttt	36960
accacatttg	tttggaatgt	ctacagaacc	tacagtttct	aaatcagaaa	ctccctaggt	37020
ttttgctatt	ttggcaagcc	attgaagtcc	ttccctctcc	ctttactacc	agaaaggtgt	37080
gtattttag	agctctctat	aatgagaaag	cactctataa	catggttgat	tcattcatttt	37140
ggagtagaaa	agtatgaatg	gaaagtccaga	gacataaaaa	taaagcccag	aggtctgagt	37200
cttagcttca	ttacagactt	tcttggggga	tggttggtta	attatctaca	cattctatct	37260
tgtctttata	attttaatag	ttaaattttt	accatgtgcc	tcaaaaccgt	tagagaatta	37320
atgagctctt	tgaaaaatgc	ttctaagt	cttgtattgc	tctaatagaa	tgctatctat	37380
gttattattt	atttctgaga	ctaaaattgt	ttacatcttt	aaactgggtg	tccttttgtg	37440
tatttttagga	tggaagtcc	agctctgata	gtgacctaa	tctgtcttca	atggtctttc	37500
tttaaccatag	cagtggctct	gacgattcag	ctggagatgg	agaatgtggc	ttagatcagt	37560
cccttggtttc	tttagagatg	tcagagatcc	ttccagatga	atcaaaaattc	aatatgtacc	37620
ggctctactt	tggaagctct	tatgaatcag	aacttggaaa	tggaacctct	cttgatttgg	37680
tatcttgtgc	agtcattcatt	atacagttct	gaaatataaa	gctatatgtt	ggtgtaaagt	37740
tgcaagtatt	tctctcctaa	ccagccccac	atattcttcc	tggttggttg	gttcttcagt	37800
aaaatagtct	tgtttcttgc	ttacactaat	tggtaat	cattccttgt	taagattttc	37860
aagacagggc	tgaggagcaag	gaaccaaagt	agcgcgtgg	tgtgattacc	tttggtttct	37920
ttgaggtttc	tcttacctag	tggttttaaa	acatcttttag	gagcagttcc	attttatagt	37980
aaactttaaat	tctgttatca	tgaacagttg	aggataatga	ataatttgat	acaataatgt	38040
aagaaattcc	tgaaaacaaa	gtgttatctg	tgatactttt	gctgcatagt	aagcacaatg	38100
aagtgtactg	ataatgtttc	aacaggaaag	tgtttttgatt	aaatgtgggc	agtatcactg	38160
ttctactagc	attcaacatc	tcttctaaaa	attaatagt	gttccactgta	attttattgg	38220
tacatgtaac	atctgtacat	gtgttttggtt	atctatatgt	ttcctgggtt	tttgtacatt	38280
tgctttatta	attttaggctt	tttttttttt	tttttttttga	gacagtctca	ctctatcatc	38340
cagactagag	tgcaagtggca	caattatggc	tcaactgagc	cttgacctcc	tggtgcttagg	38400
tgattcttcc	acctcagcct	cctgagtagc	tggaactaca	ggcacatgcc	acctgcccac	38460
gctaattttt	gtatgttttg	tagagacgag	gtttcaccat	attgcccagg	ctggtctcaa	38520
actcctgggc	tcaagctatc	tgcgtgcctt	gacctcccaa	agtgttagga	ttacaggtgt	38580
gagccactat	gcctagccta	actcagactt	taaaaatata	aaagcaattc	atttttattc	38640
ccaagaacag	taaggtggtg	gttttaatttt	agtctttaat	tctgttttta	atttattcta	38700
tttagaaatg	tcccagaaac	ttagtataac	tttactttct	gaaaatgaag	aaacctgtcc	38760

ttgggcat	gtgtgttgga	tttaagcaac	aaagttaaaa	aaacctaccc	tgtgttatgg	38820
caattttcac	ttgatgggtg	ttctataaca	caggatcag	tgaaccttta	taaaagatga	38880
acaacttttc	agcttgctta	atttcagtta	attaacatgt	atacttatct	atgttaatgt	38940
tttattgctt	aaaatgttta	atttttatat	ttggtaaaca	gatagttttt	tctctcccc	39000
tcttccttcc	atctttcatt	actacaattt	accatgcaga	gctcacaaatg	tctctctgca	39060
ccaagctcca	tgactcagga	tttgcttgga	gttcctggga	aagaccagtt	gttttcacag	39120
cctccagtgt	taaggacagc	ttcacaagag	ggttgccaaag	aacttacaca	agaacaaaaga	39180
gatcaactga	gggcagatat	cagtattatc	aaaggcagat	accggagcca	aagtggagta	39240
tggctttttc	cccctcatta	taattgttaa	aacttcttaa	aaattgtttc	acccttttga	39300
tatatatttc	tttgacttat	aaacgagcta	tattttataaa	caagggacca	gaacacatta	39360
actcagtcac	ggttatgtgc	ttccttgctt	tcaatgtttc	attatcttat	aaggaagaga	39420
acgtatggtc	tcttgaaaaa	actgacaata	agaagtaaca	actggactac	cacatttttt	39480
tttacaacct	taattttaact	cttcgtcaat	ttcttttttt	acttaaggag	gacgaatcca	39540
tgaaccaacc	aggaccaatc	aaaaccagac	tggctattag	tcagtcacat	ttaatgactg	39600
cacttgggtca	cacaagacca	tccattagt	aagatgactg	gaagaatttt	gctgagctgt	39660
aagtaacaga	ttctgttttg	gaagtacagc	tactattaca	agtgcacatg	tattacactt	39720
aaacctttta	agttcgtgtt	taaaataaaa	atattttgaa	tattttaaaag	ctaattcaaa	39780
aaatatgtgt	cgtagctatg	cattaaaaaa	ccccaaaatg	tcagaagtac	agaagtcaaa	39840
attgagtttt	cattaaccag	ttcatttgat	tatatattgaa	ttattcataa	tggtactcatt	39900
taatttttagt	aactttgggc	tgggtgctgt	ggctcatgcc	tgtaatccca	gctctttggg	39960
aggccaaggc	aggtggatca	cctgaggtca	ggagtctcag	gcaagcctaa	ccaacacggg	40020
gaaaccccat	ctctactaaa	aatacaaaaa	ttagccaggt	gtgggtggcat	gtgcctgtag	40080
tcccagctac	ttgggaggct	gagacaggag	aattgcttga	accaggagg	tggagggtgc	40140
agtgaagcga	gattgcacca	ctgcactcca	tccagcctgg	gccacagagc	gagactgtgt	40200
ctcaaaaaaa	aaaaaaaaaa	atttagtaac	ttcgaagaaa	taagaaggaa	aattaaaagt	40260
tgaagtgtat	tctaattgtat	agttttataaa	attttgttat	aaaaatacct	gttttgctt	40320
caaaataatt	tatattaata	ttttattgac	ctcaagaaca	tttaaataca	ttcagattta	40380
ttcatttggtg	gaccacattt	gttatacatt	ggattttaaag	gatccttgca	attgagttta	40440
tggccaccta	tgcatctgag	acccatggac	tgggaaccat	tctaggtcaa	tgattcagt	40500
tgatttcaatt	taagagatgt	ttattcctgg	tctttagaag	ctgctaccct	ttgttatcta	40560
attttgagct	actttgaaat	atgtatgtat	gtgtacatac	gttagtgcta	tgtatttatt	40620
aaagaagaat	cagaaaaacag	aggtaaggaa	aaataaggaa	acaaatttct	gttaagccca	40680
ccacctccca	aagcatattt	gtttatatgc	ttatatatgt	tttcctatta	tggtaagaac	40740
agtctgtaca	tattgctata	tagcagtccc	cctttatcca	catacatcct	gaaaattgtt	40800
ttacatttta	aatgttaact	actttattgt	ttttaaatgt	catttttatag	tgtagctatg	40860
ccacaataatc	caatttttag	acattttaat	tgtcccagc	caatgtggta	atgaacattc	40920
ttgcagctga	atatatgcac	atatctaatt	gtttcactag	gatagagggtg	gaattgtata	40980
acagggagct	cacatttttt	aaggcttttg	aaatgtattg	ccaaattgcc	tgccagatat	41040
actgcaccat	cactaacatt	gtgtgttgca	gtatttttct	aaacttggcc	cttttgattt	41100
tagaaaaatg	atatcaataa	tttacatttc	tttgattaaa	gtgtagaagt	tataattttt	41160
catattattc	attgtcattt	gtattttatc	ttttctaact	tgtctcttca	tcccctttgc	41220
tccgttttct	attggagtgc	aactttattt	gtaagaattc	tttttaattt	ctgtgactgg	41280
aatttttttt	tctagtgtgt	tatttcccgt	tcatcttcta	aaatataatt	gtgtttgcca	41340
acaatccatt	atcttttggt	ttgtaatgg	agtatttata	catattaaat	tatctctttc	41400
ttttttcaga	tatgaaagct	ttcaaaatcc	aaagaggaga	aaaaatcaaa	gtggaacaat	41460
gtttcgacct	ggacagaaaag	taacttttagc	ataaaatata	cttctttttg	atttggttct	41520
gttaagtgtt	ttgatggctt	ttccatatgt	tgtaacagga	aaaaaatgg	gtctatgaat	41580
ttcttcttaa	tttaacaaat	ttgggttaatt	tataaaatca	cagattggta	aatgctataa	41640
ttatgtaatg	atcaggattg	agattaatac	tgtagtataa	attgggacat	tataacagat	41700
tccatatattt	atctcctaaa	atctaaattc	agtctttaat	gaaataatat	tagccaaatg	41760
gtggaactaa	tttatttctt	ttgaggaaaa	gataataaag	aatgtaatta	aattttaaatt	41820
tcttggaatt	cccagttgta	tattcatcac	ctttgtagca	tttgacaaat	tttatgctta	41880
cgagcttctt	cactgttttg	aaataaaaata	tcctattacc	tactgatata	attatctgtt	41940
ctttgtatat	caaaaaatgt	gaaattttaca	cataaattcaa	atacatttaa	ttatccgctc	42000
aaccagaaat	gaaatcacat	ccctctacta	tactacatcc	agctccaagc	ccaagatatt	42060
taaatgacat	ccattcctct	cctagtcca	gttatgattt	tatcttgata	ttctctcata	42120
tatgaactaa	attataaagt	tagccaccat	caatacaatc	tgcgtatcta	atatcttaac	42180
tatatagtaa	tggggtaagg	gaacagcaaa	aaggagaaca	ttaattaaaa	tatacaagta	42240

agcctgggca	acatagtgag	accccatctc	ttaaaaaaaa	aattagccat	gcatgatggt	42300
atgcctctag	tcccagctac	ttgggaggct	gaggtaggag	gatcacttgc	tcccaggagg	42360
ttcaagggtc	taaaaccagca	aagctcagaa	tcccagggga	tagaaacaaa	gacttagtgg	42420
atcactagta	ttaaactgag	acacgtcacc	ctgcattgca	ctttgtttct	cagttctttg	42480
atgaaatcac	tgagctgaca	tacctgccct	cttttcacca	taaagtgagt	ttaagatgac	42540
gaagcaatgt	ctatgggata	gcctaacaaa	caatgtaaaa	accatttagt	aagttcatga	42600
aggggtgggtg	tggtaaaaaat	ttggagaaca	tacaaaacaa	atacaattcc	aaggtgtgtc	42660
ccctccagga	aggacaaatt	gctgcctgct	ctgtgataga	agaggatcag	atgtaatcaa	42720
cctgccgtca	gacttgggct	gttctctcct	gggtgtggac	ttgcctgggt	ggtcactgct	42780
gctgacaagt	aggctgtcaa	tatagctggg	ttgtcatgtc	agctgtgggt	agggggaagt	42840
ccacattgtg	gaggccacat	ccctgcactc	ttggccaatt	tgaccatgaa	tcttaagcac	42900
tggggtggct	ggaaaagaca	gccgattgac	atccatacag	aggtcatctt	gaccacttga	42960
ttagtataag	cactgaaggc	ttttaactga	gcattcacat	aggacacaaa	tattctgatt	43020
ctttggggccc	attccaagaa	ctctgggcat	acttttctct	cagacctcat	accagtttgt	43080
gttctttcca	aattttctggt	catctgggtta	tgttattagc	cactatctgt	gaatcagcat	43140
agatttttat	atcagacatc	tctacctcct	gacagaatgg	aggagatatg	ttacttaaca	43200
attctgttcc	cttgggaagat	ttcctgtctc	cactgtttgt	aagggtact	ccctcaatgt	43260
agcagtaatg	ctttcactct	gatgggaagt	cacagtggaa	ttctgggtct	ccaagaatta	43320
gtgttagtgc	atacacagtg	tctgataatc	cccagagtgt	ctgggtgccct	tggatcctgt	43380
gaagaaggct	tggagaaaag	aagattcatg	gcaagaactt	gtgatgtgat	gacagggcct	43440
tttctctggc	tcttcattct	tagtctgacc	taggtgtgag	aattagggtca	ggggccatga	43500
ctatattgtg	gtgactcaaa	ccaggccttt	gtttactaac	tgggagattt	ttacatttga	43560
agaatcaagt	aggatctttg	cccatgtatt	ttggtcttaa	gaacacaaat	gatatggctc	43620
caatgactgg	aggaacacca	gggtccttgg	tctcacgtcg	atttagataa	aacgactgtc	43680
aggcctctga	gcccaagcta	agccatcctc	ccctgtgacc	tgcacgtata	catccagatg	43740
gcctgaagta	accaaagaat	cacaaaagca	gtgaaaatgg	cctgttcctg	ccttaactga	43800
tgacattcca	ccattgtgat	ttgttcctgc	cccatcttaa	ctgagcgatt	aaccttgtga	43860
aattccttct	cctgggtcaa	aacctcccc	actgagcacc	ttgtgacccc	cgccccctgc	43920
cctaagagaa	aacccccctt	gattataatt	ttccactacc	cacccaaatc	ctataaaatg	43980
gccccacccc	tatctccctt	cgctgactcc	tttttcggac	tcagcccgcg	tgcacccagg	44040
tgaataaac	agccttggtg	ctcacacaaa	gcctgtttgg	tggactctct	tcacacggac	44100

<210> 64

<211> 16869

<212> ADN

<213> Homo sapiens

<400> 64

aagcttttagt	agagatctca	aaaatgggtg	gatggtagca	aattactaag	aactctcaaa	60
gtttctaaaag	ccttagtttc	agcttgctag	aaaacctatg	ttgagtatta	tggctagtgc	120
catagtttag	ttgggaaatg	tctttgagga	gacacttttt	cactttgtat	tcactctgtac	180
attttctggt	acttgcattc	tgtcatgctc	aggctattag	agcaggtaca	tttttataac	240
tggaaatgttt	atgtgtagtg	aagctctgag	aggactttgc	attagatctc	agcagcataa	300
tcagaagggtt	gtcctttgtc	tcagcaattt	ttaagctaatt	agtagcagaa	attgcagtgg	360
aaatagactg	ctttggcaca	acattcagaa	aatcatttat	ctttttattg	cagttcttgt	420
caccaaacaaa	tacatttttag	tacttctcaa	attgcagaac	tctcataggg	ctgggaaaaat	480
gcctgtagac	acatacatac	tatgaatgtg	ctaattgttt	ttgtattttc	atagcccatc	540
aaagctcctg	agtcagtttc	cactataatc	actgcagaat	caatcttcta	caaggtaagc	600
ttttgtagag	ttactgaagg	aagagttggg	cctagtgggt	aatgtgccac	taaaatggtg	660
gattagtcta	aaggtctctg	ctactcttta	tttgtataag	gtgtgattat	acttttttgt	720
cccttcttag	ctgttttccc	ccataagtgg	ctgttattaa	aacatctcat	ctagagctga	780
agtgaggagg	gaaagtgcct	actgacacat	gatgtgagga	tcttaagtat	tttttttttag	840
tgtagattgt	aggaattatt	cttaaaatgc	tgattgtata	gtgtggagcc	atggaagact	900
gagccgttag	tgcatgggca	ttgaagaatg	agaaggacag	agacaggatt	tggactagta	960
gaggttgctg	actgtgggtg	caaatgggta	gagtagggcc	agagattcta	aaatgccttt	1020
aagtggagtt	gagctgagta	agggcagtag	tgaggattaa	cacctactag	aaattcatag	1080
tgagaggaat	tccaagatgt	tttgataaaa	gaatgaggag	gtcagggttc	ccagggccaa	1140

agtccatgaa	catctgatac	ctcagtgaga	gaagtgcacag	attgtttgtgt	ttaaaccaga	1200
agtcttagga	aaggaattag	aacatagacc	cccaaggctc	ggcaggcctg	gcacggcaca	1260
ggcagcaacc	attgaaggct	atttggtgtt	togggatctg	aactgtcatt	taggggacag	1320
tggtgtgagt	tagtacttta	tacttgaccc	aggtggactg	agaaactcaa	gtgatgatgc	1380
ccttaagtat	actttttttt	aagcccacaa	tctatatagt	cgaagtctgt	tcctcccaac	1440
aggggtacac	tggcattcct	cagcagggt	gggaaaaaac	aacaacaaaa	aaagtctgta	1500
cacaggcaaa	catctctctt	atttttccaa	catttaatac	attgttaata	aaatatctaa	1560
agtttagcaa	acagttgctg	tgtatcagtg	gctgagcatt	ttgcatgctt	tatttcattc	1620
agttcactct	atgaggtgga	tactactatc	cccattttct	agatgagAAC	attgaggcac	1680
agcgaggtta	attaacttgt	ccaagatcac	atagccaaca	agtcatggag	tgaggcagtc	1740
tcatgccaga	gcttaagcct	agagcatagt	tcctggctct	acagcttttag	caagtgactg	1800
gctatgtgac	gaggaccaac	ctctctaagt	tctcatctgt	aaaataggaa	ttgtaaatag	1860
ttactacctc	agtgggtcaa	atgaaatcat	atgtgttaag	cacttagcag	agtaagcact	1920
caatgaatag	taggagttat	cacatcttcg	tatttgtgca	ttaccttcac	agtttacaga	1980
ttaaggccag	aagcaacttg	ttgagctacg	ggtttagtgt	actaacagtt	tccatgtgtg	2040
tctccatgga	aggggtgtgtg	ggacctgtta	ttgtgactgt	ctgtactttc	gtattgttgt	2100
ctgccaccca	tgtttattaa	atgataagga	caataatgca	acaaagtagt	caagtaatgt	2160
tatgtgtaca	cagctattgta	gtggctatca	cagcagtgcc	actggcaggc	agcaccatgg	2220
tggcaagttc	aagaggtcac	tgccagccac	tgagctagag	cccagatcag	gcatgcaaga	2280
ggagcctgag	tgggagccac	tggggatcac	ggccaagagt	gtgaccaccc	aagaccaga	2340
atggctgagt	ggcctccctg	gagcatggca	gtggcagaac	aactccatga	actcagatct	2400
ggtgatgcct	aaactagtgc	tgttctcgtg	tggaccctt	ttctctacca	gaaaccttga	2460
atcctctcag	caaatgagga	gactactcag	atcagtgact	tagtctgtt	tggtgttata	2520
tatgtgtaca	caacacagca	catattaata	aatacctact	atgtgccagg	cactgcctac	2580
cactggaatc	tttactaag	acattgtttt	tactttgcat	ttctgccttt	acactatgaa	2640
agtagatgtt	ttggattcat	attcattcag	catacatttg	aatatgctgt	gttatgcata	2700
gtaagcctat	gataagcaag	tattctcatt	tagaatttgg	gaatattgat	tatacatgtg	2760
gacaaacaaa	ccataaatgc	aaactattta	tatgataaat	aactttggac	tgatggctgg	2820
gaggaaggac	cagctattga	tgggtaggaa	ctagcaagta	gcggactgtg	gcctgcatac	2880
accagaccca	tccgtagtga	tcagatgaa	acagccacc	tcagacactt	ggataaagg	2940
tccaccagga	aaaaactcct	ggcctatcag	gtgctatgtt	acagttcagt	tactggaagt	3000
atttccctcaa	aagtgttttt	atggttgagg	tacacattcc	tacagcttta	cctgctgcca	3060
agtccctggt	tcaagggaag	cagcaatgaa	ttacactgtt	cccgtagtca	aggacagtat	3120
atcttaccaa	gaactatacc	cacttaagga	ggtgctggat	gtcataaaga	tttggatcaa	3180
ccattatggg	tgttcagagg	agagattatt	tccagctcaa	gaccagggga	agaggacata	3240
ggatggatac	cagagtcata	gggaggattt	acacaggac	atgtacacat	tagttagttg	3300
ggtataaagt	ggaacagaaa	tgaatgagac	acaaagcctt	gaatgccaga	aatactagta	3360
gtcctgttgt	ggaaggatat	aaaactcaac	tgggagtggg	agagaaaggc	agcagtgaat	3420
ctaggagatg	tacagtaggt	tgaggtaaac	atatcctgaa	gactataatc	caaagattat	3480
ttttggtttg	aatttggttt	ggtttgaatt	catggtatct	attttctttg	agtggatggc	3540
tggggagggt	ggcatgtaga	atgcattctt	accaaatcag	catgattttc	aagacagtac	3600
agagaaaaga	ctgctgagct	gatgtaggag	ctttggctgc	agtctctatg	gctttcagca	3660
agccgtttta	ccttactact	gcttcatgac	tgtggctaac	aaagtaggga	tagtacggag	3720
cacagaggat	ttttagggcg	gtgaaactat	taatactctc	tttgatgat	actataatgg	3780
tgggtacatg	tcattataca	tttgcccaac	cccacagaat	acacagcacc	aagagtgaac	3840
cctaattgtga	actctggtct	ttgatgatgc	tatgtcagtg	tacgttcac	cgtgtaacaa	3900
gtgtaccact	ctagtgggtg	gaggggttat	tgataatagg	ggaggatgtg	catgtgtggg	3960
ggcaggaagt	atatgggaaa	tctctctact	tctgctcaat	tttgctgtaa	acctaaaacc	4020
tctgtaaaaa	ataaagtcta	ttttttaaaa	agtggggatg	gtattacggc	aatataaaat	4080
caaaatactt	tatgaacaaa	tcttttctcc	agatgtaaac	tgtcatatat	gcacctcgt	4140
atgtgtatgt	ataattttca	ttcaaactgt	aaacaacttt	agaattggca	ccaaacatat	4200
aaacactgat	acattagact	atctcgaaca	ccttttactg	accactttga	aaacttgctt	4260
acctattaag	gttcattcat	agctgtgatg	ttctattttt	attttcaatg	tgggattatc	4320
ttctgtttcc	cccaggaggt	atattaccaa	attggtgatg	ttgtttctgt	gattgatgaa	4380
caagatggaa	agccctacta	tgtcacaatc	agaggtttta	tccaggacca	gtattgcgag	4440
aagagtgcag	cactgaactg	gctcattcct	acctctctta	gccccagaga	ccaatttgat	4500
ccgcctcct	atatcatagg	taagtttgac	aatggcaca	ggtttttttt	taacttagtt	4560
aactctccaa	tattatgtaa	aagagtgtgt	tagtcagctt	gggctgtcag	gacaaaaatat	4620

cacagactga	gtggcttaaa	caacagaaa	tcactttctc	acagttgtgg	aggctgaagt	4680
ccaacatcaa	ggtgctggca	acacggattt	ctggggaggg	ttttcttcct	ggcatataga	4740
tggtcacctt	cttgctgtgt	cctcacatgg	cctttcatgg	agtgagagct	ctttgggtga	4800
tcttcttata	aggacacccat	ttctgtcaga	tgagggcccc	acccttatgg	tttcatttaa	4860
ccttaattgc	ctccctaaag	gtctcatctc	caagtaccat	cacattgggg	attagggctt	4920
caacatataa	atltggaggg	tggcgggggg	ggatgcaatt	cagtccataa	caaaaaaagc	4980
atgagtatta	ttaagtacaa	aaaaattaga	gagctttata	gaaaatatga	ggcattttat	5040
gtagctggag	tgtgagtgtc	atcagttatt	ttgagttaga	gcaatgtgca	tctactaaga	5100
agtggtatgg	ataagatttt	tttggagtga	cccagggtta	aactgtacta	caagaatgta	5160
ttgctcagga	actaggttat	ttaggttact	tatttatata	aacctattca	aaaataattt	5220
aggaaagaac	tatcccagtt	atcccatact	tgcaaattct	caatatgtgt	gacctgtcat	5280
gctacacatg	tcatcttagg	cctttatagt	ataaaggctg	atagttgaaa	tggcagctgc	5340
tgtgcttttg	ttaatttcaa	agctgccaaa	acagttgtga	gatagactca	caagaattta	5400
ctgattaata	caatttttaa	agttttcaga	tttttacagt	tacttcagac	tttttatctt	5460
tctgcagtga	gcatgcatca	ttacttttgc	atcctgagaa	caagcataag	tgtgtttttg	5520
gagagaactc	cagggacaaa	taatatacca	ctgttattct	cacctatatg	tcaagtttga	5580
tacattacca	aacaattcta	gocctctgct	tataagtata	tagaattttt	atttacctta	5640
tctatggatc	aggatctcag	cagaggcagt	gatgtatcag	aatcaccttc	gggattcctc	5700
tactgcctcc	tctttctaat	ccccagattc	tgatatgcat	ccttgtccta	cagcgaggca	5760
gcatggcatg	aggtcagaac	accagttctg	gagccagact	gtctaggttc	acagcctgcc	5820
atltacoggc	catgtgactt	tggcaagttt	ccttagtctct	cctgcctcac	tttccctcata	5880
tgtaaaatgg	gaataataat	agtgcctacc	tcagaagggt	gatgtgagga	atgaagggtat	5940
tgatacatgt	aaacttagag	cagtgtgggt	acaaaataaa	catgatgcaa	gtgttcaatc	6000
actgtttttg	ggagaatgcc	atattcttta	agccgttaaa	gaagaaaaaa	tgattaagaa	6060
taattttcaa	gtaatgcatg	tttcaagggc	taatgccagg	ttgctcccag	agtggctctct	6120
cccagtgctc	agaaatttta	acatcttatg	aaaatgatat	atatgggtcaa	aatgtatttt	6180
aacctttccc	ttggctgcct	tccagggcca	gaggaagatc	ttccaaggaa	gatggaatac	6240
ttggaatttg	tttgtcatgc	accttctgag	tatttcaagt	cacggctcatc	acctttccc	6300
acagttccca	cagaccaga	gaagggtac	tatttgactc	atgttggggc	tactcctga	6360
ataacaatta	aggaatcagt	tgccaacctat	ttgtagtcca	caaattaaaa	ctgggtttcc	6420
aggcctgggtg	tgggtggctca	cgctgttagc	cccagctatt	gcaccactgc	tctccaagct	6480
gggcaatgga	gtcagattct	ctttcttaaa	aaaccacaaa	aaaactggat	ttccagttct	6540
ctaataattct	tagtaccaca	agatatgtca	taggtatctt	taaatgaaat	tcttagctgg	6600
aaaagtgact	aaaaagtttt	tctcctgcta	cctagtaata	aacaaatcat	tgtttattac	6660
tggtcactta	gaaaattaaa	agggataggg	ccaggcacag	tggcttatgc	cttattattgc	6720
agcactttta	gaggccgagg	caggcggtatc	acctgagggtc	gggaagtggg	tcgcctgagg	6780
tcaggagttc	gagaccagcc	tggccaacat	ggcgaaaacc	cgtcgctact	aaaaatacaa	6840
aaattagcca	ggtgtggtgg	catgtgcctg	taatcccagc	tatttgggag	gctgaggcag	6900
gagaatcgcc	taaacccagg	aggtggagggt	tgtagtggagc	caagattgca	ccgctgtgct	6960
ccagcctggg	caacagagtg	agactcttgt	ctcggaaaaa	aaaaaaaaaa	aaaaaggctg	7020
ggcacagtgg	ctcacgcctt	taatcccagc	actttgggag	gctgaggcag	atggatcgcc	7080
tgagggttggg	agttcgagac	cagcctggcc	agcatggtga	aacctgtctc	ctactaaaaa	7140
tacaaaaatt	agccagggtgt	ggtggcgcac	acctgtagtc	ccagctactc	gggaggctga	7200
ggcaggagaa	ttggttgaa	ccaggaggcg	gaggttgtag	tgagcagaga	togtgccact	7260
gcaactccagc	ctgggtggac	agagcaagac	tccgtctcaa	agaaacaaac	aaaaaattaa	7320
aagggataga	atataatgaa	atatattttg	aacttaaat	atattctata	tgtgtatctt	7380
cctaggcaaa	agctgtaatt	tccagagaga	ccattaggaa	caggtagtat	ctatttttct	7440
ccattatttta	tttctagaaa	ctcataaaat	ggattgtatt	tttctataag	aacaaaatat	7500
taattaagggt	atagatgact	gaccaagggc	ttaatcaaat	aaaatgacta	acagcatcta	7560
tcataaagcc	acacaagcct	tatgttctca	tctcaaaaaat	gctgtgacag	ctttttggct	7620
gctttaacca	taagaaaaat	gattggtgga	tgattttatt	agcccaggct	tttaaaaaat	7680
ttcatctagg	ccacgtgcgg	tggctcatgc	ctgtaattcc	ggcactttgg	gggacctgag	7740
tggatggatc	acttgaggtc	aggagtccag	gaccagcctg	gccaacatga	tgaacccttg	7800
tctctactaa	atatacaaaa	attagttggg	tgttatggtg	catgcctgta	atcccagcta	7860
ctcgggaggc	tgaggcagga	gaattgcttg	aactcgggag	gtggagattg	cagtaagccg	7920
agatcgtgoc	actgcactcc	agcctgggtg	atagagcaag	actgtctcaa	aaaagaaaaa	7980
aaagaaaaaa	ttttaatttta	atccttctgt	agaaacaggc	attcagaacc	attccattga	8040
tcttaataaa	gctgctcttt	actgtttcta	gtcaaaaaatg	agacttcgat	caaaccataa	8100

gattttatac	tgcagatagt	cagcttcacc	aaagccgcag	aggaaacatg	tcgagatcag	8160
gcttctctgct	tgatagtcctc	ttgactacca	ttaaaaacgaa	tattggggag	tcagtaaagt	8220
cattggtagg	ccattagcat	tgatatcttt	aaaacatcta	ccctaaacca	tctgctatgg	8280
acccataata	agaggcctgt	tgtatatgaa	attgtctaga	attcaggtgc	aggctcttgc	8340
cggtttaagta	agggagcaac	acgtaaaaatg	ggagaggagt	ggggtgtact	cacttgcctc	8400
ctcttttgtc	ctgatttaac	cagcattttt	caaccctggg	aaaattttgca	gaatctaagt	8460
tgattgtaat	gatttttgagc	tgacgcagct	ttactctta	ccctttttcc	acatagttat	8520
ggtgtttgag	ttggaaaagaa	acaactatag	gtagctacac	gtacataatt	atctctttat	8580
tcacaaaagg	tatagtaaaa	ttgattgtaa	ataactttct	aagtgccaat	attcaaaact	8640
tttgatttaa	aatgtatttt	tcaccgtgca	tttacttttg	atgtattttat	ttcattttaa	8700
caattttaaat	ggggctcttt	aaccaaaaat	ggtattttaaa	accaaaaacag	tatcgctact	8760
agaattttgga	gtagaggccg	ggcacagtgg	ctcacgcctg	taatcccagc	actttggaag	8820
gctgaggcag	gcggatcacc	tgaggtcagg	agttcgagac	cagcctggtc	aacatgaaac	8880
cccgtctcta	ctaaaaatac	aaaaattagc	tgggcgtggt	ggcgtgcgcc	tataatccca	8940
gctagtctac	tcgggaggct	gaggcaggag	aatcgctgga	actcaggagg	cagagactgc	9000
agtgaagcca	gatcgcgcca	ctgactcca	gtctgggtga	cggcatgact	ccatctccaa	9060
aaaaaataat	aaaagatttt	ggagttagatt	catcattaat	aagtaacaga	ttttaggaaa	9120
atcaaaaaat	ggctaataaaa	atgaacacaa	tgtaaaacat	ttattaaaaat	gtagactttt	9180
aaaaatctat	aaattgatca	tctgtttata	aattggcaga	tggttggtga	ccatctttta	9240
aaataaagat	tgaattttcac	ccagtgtgat	ggttcccatt	gcttatattt	ctcctgctga	9300
ggccggacct	gatattggccc	tggtctgtgt	tcccagcctt	gtttcctcat	taccactaaa	9360
atctttcccc	tgtatgccc	cccaattttt	ctggctctga	gtccttggtc	atactgttct	9420
ctccaattct	accttccaaa	ggcctttctt	aacaccttcg	gattccttct	ttgagaactt	9480
tccagattcc	catgcctttt	tggaatcaat	ctctatccta	ttgtcatcac	atttaagttt	9540
ctacttccat	catcctcact	cctatccctt	tggtcctggg	atgacaggga	tgctgtgttt	9600
tatttactca	tctttgtaac	ttccacataa	cctaaccctg	gttcttgctt	atgggagatg	9660
ctgattgtag	ggtctgagtt	agatactgtt	aactaaaaatg	cttggtgata	ttttagttat	9720
taattcatat	taactttggc	tgaactttt	aaattctatt	gtgaatagtc	aagtaaaatt	9780
tagattgtta	cattctgggt	tagtattaga	ttgtttttta	gattgtttta	aacaagatgt	9840
ttttaagatg	agtttttaaat	agttctctta	acacaaaataa	agcttaatat	gagtatttga	9900
aggaaattat	cccaaaccat	tccagttcct	ggctgtgaaa	ggcttttcca	ggcctaataa	9960
gttttccact	tcagccgtaa	gtaggtgaaa	tcaaatgaac	aatagaggga	aatgtattta	10020
tttgctttat	acacatgcat	gtgtgtgtgt	tctacatata	aacattgcac	acgcttagaa	10080
tgaagtttct	gtcatgccc	gaaaaggag	aggcattttt	gtggattttg	tctggctgcc	10140
ctggggatgt	ttgaagaact	gtgctgttta	cttcatacca	gggtgtgtgag	ccataccttt	10200
ggtaggaggg	tatacctcct	acacccaaga	aatataagcc	aggagaaggt	ctgtgccaa	10260
agaaggaaac	caaagagccc	acaagaggtg	ggccattaat	tattgggtca	gatgcataaa	10320
tgacagtaga	tttattttaag	cacctcttaa	tggtgaccca	caagggaagat	tgctcgtagt	10380
agcggaaaag	ttcacaataa	ataagagaaa	aaagcagaat	gtagaactgt	atgatagcaa	10440
ttctgcaaac	aagaagcatc	ttttataaaa	gatggaagga	gcccaggcac	agtagctcat	10500
gcctgtaatc	ccagcacttt	aagaggctga	ggtggaggat	cacttgagct	gcagtgaacc	10560
atgattgtgc	caccactcca	gcctgggtga	tagaagttag	accttctctc	aaaaaaaaaa	10620
aaaaaaaaaa	aaagacggaa	attcctccag	aatttttaaca	tgtcaacaga	ggttttctgc	10680
agctactttt	ttcagcttta	tacttcgcag	tattttccaa	attttctcta	acaagcagta	10740
ttttccaaat	tttttacaat	aagcacacac	acacacacac	gtttgtttgc	ataagtgcc	10800
aactggtggt	gaacaaccgc	tggctttttg	tctatacata	tctagaatat	tttataaata	10860
gtagttctta	aaccttgtaa	aggagtgtaa	tgaccagctg	agaaaataaa	gtcagtgatt	10920
tcattatttt	cctatatatt	catcatgatt	ctaggaaaga	acttgggagt	gacttccttc	10980
agcttcagcc	actcctgggc	caggcgcatg	cttagctctg	tggtaaaagg	caccagcttc	11040
ttctgcaggg	tgctgtatc	atctgaattg	gaggtttggc	gagggtaaga	gactgatgta	11100
ggttcaagtt	tttctttcct	gtcctccact	tgaaatctgt	cttcccttcc	agactgcctg	11160
cgctgtgtac	tttaaggcccc	aacaccaaac	acagaagcaa	cagccttaca	cagagtgttc	11220
agcaagctcc	aacaattgtg	taaggtaaa	tttcttttat	agattccttt	tctatatcgc	11280
tcctagtgg	tctgtttctc	tgatcgaatt	ctggctgata	acagttgctg	agactctgaa	11340
agagaaggca	aggaactact	gtttctcatt	ataaactgtt	tagaattatt	tggccatctt	11400
tttgctatga	atatgtagt	ctttgatata	ttttttaaat	caaaaagtaa	tgaaaagagat	11460
cacataggga	aagatagatt	ggattatttt	taaagtttat	ataactaaat	gaaaagcaaa	11520
gaataaaatg	ggagaaacag	ctccctcatg	tggctgttgg	caggaaagctt	ccattcctct	11580

ctgtgggcct	ccacaggttt	gctcacagca	aatgggtccgt	gacagaaaaga	cgcaagggca	11640
gttgcaacca	agatggaagc	caccatcttt	totataacct	aatctgaaag	aagggacata	11700
ccagcacttc	tgccatatgc	tggtgggtca	cacagaccaa	ctctggtaca	gtgtgaacac	11760
aggaccacac	aagggcgtga	attccaaggg	cagagaccac	tagggaccac	ctcagaggca	11820
cagagggaca	ccctatccag	ctgggtggcca	atgtaaaatta	acatagcttt	ttagaatagc	11880
aatatgtatc	tataatctta	aaagtattaa	aagtacttct	tgatccagta	atttcatttc	11940
taagaatcca	tgctaagagg	attttaaagt	tggacaaaaa	aatgggtata	aaaagaagtt	12000
gttaacagta	tttaaagttg	tgaaaaacca	gaaacaatct	aaaggtccaa	caataggaaa	12060
atgaattttg	atatttttct	aatagaattt	tatgctgtca	tcagaaatac	catttacaaa	12120
taatttttaa	taacgcaaaa	aaaagtttat	aaaatgttta	gtgtaaaacc	tggacacaa	12180
tacataatga	ttctgatttt	gtaaaaaaa	aaaacaaaaa	cacacacata	tacacatgca	12240
tacatatgca	tataaagaaa	actggaacaa	acaaaataac	aagcatagtt	ggaattacag	12300
tcattttaat	attcttttat	cttttaaaaa	ttttgaagtt	tgtattacta	gcatccacta	12360
cttacgtagt	caggaaaaaa	atacaacttt	aaaatagata	tttaggtcca	aagatggtaa	12420
tctaaatggg	gttacaggct	gaatgtgtgc	ctgatcccca	tgccccaagt	tcatatgtta	12480
aagccctggc	ccccaaggca	atgggtattag	gggagtaggg	cctttgggag	gtaatcagat	12540
ttctacgagg	tcacaggggt	ggagcccgca	tagtggaatt	agtgtccttt	taggaagagg	12600
agaacagacc	aaagccttcc	tttctctcct	cactatgtaa	gaagacagcc	agaaggtggc	12660
cacagccagg	aagagagctc	tcaccagaac	ccaaatctgc	tagcaccttg	ctcttgggtt	12720
ctcagcatcc	agaactgtga	gaaatgaatg	tgtgttgttt	aaaccactca	ggctacggta	12780
ttttgttgca	gcagcccaag	ctgacagaga	tagaaacaac	acaaggaccc	atcagcagac	12840
gaatggatga	tcaaaacgtg	gtgaggtcgt	gcagtgggat	attattcagc	cgtagaagga	12900
atgaaattct	gatacatgct	ataatgatga	accttgaaaa	catgttaatg	gaaataagcc	12960
aaacttaaaa	ggacaaatat	tgtataattc	cacttatatg	agttagttag	ctagaatagg	13020
caaattatgt	catagataca	gaacattaga	ggttaccagg	gttgtgggaa	gaggggtatt	13080
gtgggtacaa	attttcgggt	tggagtgtat	ttgaaaaaat	tctggaaatg	ggtagtgaca	13140
gtagtcaaca	tgatgaatgt	acttaatgac	actaaattgt	acacttaaaa	atgggttaata	13200
ctgggctggc	gcagtggctc	atgggtgtaa	atcccgaaac	tttgggaggc	caagacaggc	13260
ggatcatgag	tcacagagat	tgagaccatt	ctgggtaaca	tggtgaaacc	ctgtctctac	13320
taaaaataaa	aaacaaataa	aaaaaaaatt	agccgggcat	gggtggcaggc	acctgtagtc	13380
ccagctactc	gggaggctga	ggcaggagaa	tgggtgtgacc	tgggagtcgg	agcttgcagt	13440
gagctgagat	cgcgccactg	cactccagcc	tgggcaacag	agccagattc	cgtctcaaaa	13500
aaaaaaaaaa	aaaggttgat	acctgggtgc	gggtggctcat	gcctgttaatt	tcagcacttt	13560
gggaggccaa	ggcaggcaga	tcagttgagg	tcaagagtta	aggaccagcc	tggccaacct	13620
ggcgaaaccc	catctctatt	aaaaatacaa	aaatttagtcg	agtgtgggtg	tgggtgcctg	13680
tagtcccagc	tgctgggagg	atgaggccta	ggaattgctt	gaaccagga	ggcagagggt	13740
gcagtgagtt	gagattgcgc	cactgcactc	cagcctgggg	gacagagcga	gacttagtct	13800
caaaaaaaag	gttaaaattg	taagttttgt	tatgcataat	ttaccataat	cttttaaaaa	13860
tagatatata	ggagataaa	tcaacagaat	ttaataacca	gttgtaaata	gagactgagt	13920
gaggaggatg	aattaaggaa	gacattgagt	acaacttttt	ggtaggtgaa	aaactcttaa	13980
aaaaatacgt	gggcaaagat	cctacttgat	tcttataatt	taaaaatctc	ccagtttagta	14040
aacaaggcta	gggtggagatt	tgcatgtgat	gtgaggtgtg	tgttctgttt	tgtaatgtga	14100
ggactgtgag	ccatctcctg	gacttgaata	tccattagat	aattgaaaat	acggatttga	14160
gaactcagga	gacgtgcaat	gcagtaacaa	aactctgcac	ctagttgatt	tctgtctcct	14220
aatttaatgc	ttttatggga	caaactgtta	ggcaggtggg	caagatggac	agccatattt	14280
ttgtgggttt	ctggcctgtg	ggccagcctc	agtgtcactc	ctgaggtcat	gtccaaactt	14340
agaacacatt	caggcctacc	acagtcaagg	ctccctttct	caactctagt	cctctgcaca	14400
aatatccgaa	gcctagaaat	aataatcatc	tgtccttgtg	tcttgcatta	tgaaagccta	14460
ggaaagggcc	ttgggaatta	agaagaatgg	aaaaactggg	ctaactgctg	catgcttcag	14520
cttgacgggg	aatcactgaa	atggggacag	gccataaaa	gacaaccaga	agagtggctt	14580
cagcaaaggc	atcgtttttc	agagcaagct	agagaatcct	gccagcgtcc	tcaggcaggg	14640
ccctggggca	cagaggttag	gcaaggaggt	gtcccagcat	gttgatgccc	tgagcatcag	14700
aataatgccca	tagaggagct	tccaaagagt	tcatttcagg	ttttgtaagc	cgaacatttc	14760
taggcaaata	aaatttgatt	ttgtgaataa	agcttgtttc	ttcaactcca	gtgcagattc	14820
tcatagattg	atagtggctt	gtgatccaga	taaagaaaac	aatttttcaa	agattcatat	14880
tctttgtaga	tgtacggatt	tagagaccat	ctaactaac	tccctcattc	tacagatagg	14940
aaaaatgagg	cctaaagaag	ttaagaaaat	accatggaaa	tgtcactgct	gaactgccat	15000
acgtaggatc	cgaaagaaat	tgggtaaatg	ctactgtgag	aaatacagta	ctaggtccaa	15060

agaatctaata	acaaattaaa	aatctaaatg	ttattttctaa	agcatccctg	cacatggctg	15120
aacttacata	gtttcatttt	ctttcttttc	tgttgaagaa	gaggcaattg	gctgggtgca	15180
gtggctcatg	cctgtaatcc	tggcactttg	agaggccgag	gcgggtggat	cacctgaggt	15240
caggagtttg	agaccagcct	ggccaacatg	gtgaaacccc	atctctacta	aaaatacaaa	15300
aattagctgg	ctgtggtggc	cgctgcctgt	aatcccagct	actccagagg	ctgaggcagg	15360
agaattactt	gaatctggga	ggtggagggt	gcagtgaagg	aagatcacgc	cattgcactc	15420
tagcctggat	gacaagaggg	aaactccatc	tcaaaaaaaa	aaagaaaaaa	agcaatcact	15480
aacctgtgtt	gtttattaaa	catgacagac	tggcatgaag	taattaccaa	actgtaaaaa	15540
aaaaagctac	aatctgccag	gcatgggtgg	tcatgcctgt	aatcccccac	cttggggagg	15600
cagggtgggg	gatcacctga	ggcctggagt	tcaagactag	cctgggtcaac	atgggtgaaac	15660
ctcgtctcta	ctaaaaatac	aaaaattagc	cggcgctggg	ggcacatccc	tgtaatcccc	15720
gttactcagg	aggctgaggc	aggagaatca	cttgaacctg	ggcagtgggg	aggttgcaagt	15780
gagccaagat	cgcaccgttg	tactccagtc	tgggcccaga	gagtgaagct	cgggtctcaaa	15840
aaaaagaaaa	aagaaaagct	acaaccttaa	tctcaacttc	tcataacatc	atctctactt	15900
ctgattagaa	gagtggaggt	ggggagggtt	attacaaaaa	gactgttata	ccttacacac	15960
ttctcccat	gaatagtga	ggtgtgagt	aaaaagacag	caatttttatt	ttttttttga	16020
aacaggttct	tgactgtcoa	cccggtgtgg	agtgcactgt	tgtgatcact	gctcactgca	16080
gcctccacct	cccaggctca	agtgatcctc	ctacctcagc	ctcctgagta	gctggggacca	16140
cagttgtgca	ctaccatgcc	cagctatttt	tttttaagag	atgggggtctc	actatattgc	16200
ttaggctagt	tctcaaactc	ctggcctcaa	gcagtcctcc	gaccttggcc	tcccaaaggg	16260
ttgtgattac	aggcataagc	caccacaccc	agccagcagt	tttagaataa	aggggtgaagg	16320
tgtgtttggg	gaaatataat	ttaaaaaaca	aaatcttctc	tcaaccagaa	aatcctctcc	16380
atgaaggcag	tagagaaaga	taagctttat	tattgaataa	aaattaaatg	agaatgtgat	16440
gcacatcaca	ggcactttgc	taagagatca	caaagacaga	aggaaaatttc	accattttgt	16500
acagccaagc	aggtacagcc	cattacatgt	atgttttctga	gataaatagt	cctcaactaa	16560
gagaacttga	cagcaccact	ggtcacacag	ttcatttctaa	ctttacctga	taattgatgt	16620
gaccacttgt	gttatctaag	atatcaactt	ttcgggggtg	ggggagtggt	gaaacaggag	16680
ttacttttat	agcttgggtc	aaggtaactc	ttaagattag	gctgtttacc	tcccacagaa	16740
actggaagat	aggtatgcta	tctggtaatg	tttacttttc	ccagatcctt	gagaaagaca	16800
ttcctaggct	ataaagctga	caaaaggctg	attcagtttt	taaatatata	tatctgtata	16860
tgtattttca						16869

<210> 65

<211> 15000

<212> ADN

<213> Homo sapiens

<400> 65

gatctcttga	tcccaggagg	tcaaggctgc	aattgagctaa	gatcaagcca	ctgcattcca	60
gcctgagtga	tagtgggaga	ccttgtcttt	aaaacacaca	cacacacaca	cacacacacg	120
agggcctttg	accactcttg	agtagaagac	tcgagaagaa	caaagtagaa	ggccagagaa	180
gaacaaagtt	acttgaaaaga	tctcttatta	aagagaatgt	acaagctatg	aaaaaaaaaa	240
aacacacaca	cacacacaaa	cctcatctgg	aatgaaaaaa	acataatgca	tttggtttct	300
ggttccttag	gctgttatgg	aacaaccaa	gaacattatt	ttggtttctg	aggtcagaac	360
tattttattc	ccctcaagca	cactatgctt	atggtttgag	ggagaatgag	aaataggaaa	420
ctaggaacag	gctgaaatgg	tctaactctg	accatctaat	tctgcagtgt	cttattctca	480
ttctaaaaga	gaatggttat	attcgctgtt	ctagcataaa	aagtaatgat	aaaaataaaa	540
gatcccgtat	taccagacaa	taatccccta	gactgtttta	atgcttgggt	gagtatttgc	600
ttatgatctc	agactttaaa	agatggtctc	cccctatggg	gaagcttggt	aattatgtag	660
gcatccttaa	tgtctgttta	cttatcaaaa	ttttatcatt	gttagttgta	ttactacttg	720
acagtcctaa	ttatttaatt	gaaaagattg	gttaacattt	tatagtcaaa	gtaatgtgtt	780
cctgtgtttt	ttcctgttta	ggttatttga	gtgatgagta	aagaatacat	accaaagggc	840
acacgttttg	gaccccta	aggtgaaatc	tacaccaatg	acacagttcc	taagaacgcc	900
aacaggaaat	atttttggag	ggtaagtaag	ggaaatttct	tcagacccat	taaatgttag	960
gaaaaaatgg	agctaaaaga	gctgggtggc	tcacotttct	catcctgtgc	tgagaaatgc	1020
tggggctcac	ccataagtat	ccagcatccc	catggacaca	gggaattctg	aacaaatgtg	1080
atgaaaccga	tgaaatgtct	ggcctgtagg	tgggttagtga	tgagataacg	ggctatatgt	1140

gaatcttgat	ttttgcaatt	cattagagct	ttgtaatgaa	aggaaacagt	ttgttgcttg	1200
ctttaaggat	aggttcattt	gcatttctcc	gcaaggaagt	agtaatgagt	taccaagcct	1260
tagatttcac	coctttttga	tttcttgctg	acttaacttt	aattgaatgg	aagagttatc	1320
acaaatgaat	tatctttttg	gttttttttt	ttttgagatg	gagtctcact	ctgtcaccag	1380
gctggagtg	aatggcatga	tctcggtcca	ctgcaacctc	cgcctcccag	gttcaagcaa	1440
ttgtcctgcc	tcagcctccc	gagtagctgg	gactaagggtg	cgcgcaccca	tgcccagtta	1500
atttttgtat	ttttagtaga	gacgggggtc	cactatgttg	gccatgatgg	tctcgatctc	1560
tggacctcgt	gacccgcca	ccttggcctc	ccaaagtgtc	ggaattacag	gcaagagcca	1620
ccgcgcccag	ccaggaatga	caaatgaatt	accttataag	taaatgccat	taagggaagg	1680
tagctggaag	atgggttgag	gggaatggag	gaccacagaa	ctagtcctat	ttaaatacat	1740
gtgcatggta	aatgatttcc	atgtgacaat	aggttaatta	tctcatagca	taaggaaaat	1800
gcttaacagt	catatgcaag	atgataagct	ttcctatagc	atccaaccaa	aagatctagc	1860
cagtacaatt	tcctttgcta	tattaggggtt	agaaaggccc	ccagaggtga	accaattaga	1920
tggaaatcctt	gaataaaaca	ctggattagc	agtgaacaga	aaaaagtcag	attgctttcc	1980
ttcttcccat	agatgtctca	gggatattta	gtttcctcag	aagataaaga	atttagtaag	2040
cgtttttttg	tgcatactta	catgaaatgt	acattatttg	aattctttta	aaagaaacag	2100
ctgcatagata	acaaaaattg	tgttatgctt	gcttttagctg	gtattttttgc	ctagaacgat	2160
tatatcgttc	ggacaagaag	ctatttctaa	gaacaatat	ttttaatcca	ggaagttttt	2220
cattttttaga	aattttatctt	actatttccc	aagcaaaaaga	gggtagttac	agatttacta	2280
agaatcatgt	gctcacaatt	tttatttaat	aattattcct	ccttaaaata	tattaatcac	2340
ctgacttaca	atggttgaac	catgagtcca	tttttgcctt	tattgtcaat	aacgtcttct	2400
cagaagtga	ccacaaaggt	gcatagttct	tggagttaaa	ggtctgaatt	aagacaatcc	2460
agcataagtc	tcattaatgt	gtgatttatt	tgagaaaagg	caagaagtac	ctaagaatct	2520
ccccctcact	gtccagttcc	ctgttttcatt	taaagattca	ctgtaagtaa	ctgaaaggct	2580
ttccttgga	ggattttattt	gaatcagttc	ttcacatgca	aaggatattg	tagaacatct	2640
cgtttttgct	ggcaggaata	tgaacatctg	ttgtgaggaa	agaaaaagtt	tcatgcaaat	2700
tacactgcca	aagaagggat	gttcaagttg	agaaaccagt	gacatttctt	gtaactgtac	2760
tatgaatcag	cgcattttta	tcttctagat	aatatatgga	agtgcaggaa	ggtggttaga	2820
aacggtgttc	attttacata	tgcgttattt	tattctgtgt	gagtgacttc	ctaggaccga	2880
cattgctgtt	tttaaatgag	gatacagtaa	attgcagttc	gaggaaggct	aactggaatc	2940
aacatacccg	tagcttttaga	aagcagtttc	cgcaccagcg	aagagtacaa	gagcgatgga	3000
accccatggt	cctggaagtt	tgcacatcag	agtaaacaaa	cttgaaaacc	cctcttgata	3060
gcagaattca	cccagccttg	ttccattttc	tcttaacaaa	acacaccgca	aaagctctca	3120
caagctgctt	tgatgaagcc	acatgtattt	cccccttcac	aattttacagg	aggttactct	3180
taaaagaaag	tgattctggt	gtttaccgcc	tgtgttaaa	ggacagagtt	cctttttatt	3240
tctgataacg	tttgagcgaa	atacagaaac	tatctgtaga	ctagcatagt	cggtagctga	3300
gtaaggaaaa	gcaataacct	gctgtccggt	gagcacaaaa	ttctgtctac	gaacagtgcc	3360
ttactgctgc	ttggagactg	caagtccgag	atcacactag	gtattgactg	attgtataag	3420
gaaattttctt	aaagtctaaa	gtaaagggtg	tacctcttaa	aaagagggga	agagagaaaa	3480
ctttgtgtgg	aaggataagg	agtgtgttta	tagtttcagt	aagagtgtac	gttttaattt	3540
ttcttcttcc	tctgcctctt	tgccaagtag	cctgagtcca	tctgttatcc	agaagtagta	3600
ttactctagg	acaaaacttca	aattcttcat	tctgcgttgc	ctttaaggaa	caacatactt	3660
tcttctctgtt	cttttttccaa	aaacacacgc	ctatggctct	gtgtgtggtg	ttttagccag	3720
cctcctccca	gataaggggt	tcccttccct	cctttgcatt	gaaaggaaag	tgcaagtctg	3780
gacatgttta	tcaagaggaa	aagtgaactc	tcagtaatat	actgtcaaat	tcgggctgct	3840
gcccagagtgt	tcgcttttgtt	atggcagggtg	aagttcacct	ttgccccacc	cagtgtttcc	3900
acaaaaaggc	aaggttccaa	gtattcatat	gaacaagtgt	tacttttagga	cctggagggt	3960
tgggggtgga	ggatgtttgc	atagttgaag	ccttgggagg	gggtgttagga	aacggcgagt	4020
acagaggcca	tagaaaaagc	taagactcag	tttgacgtcg	tcagccggct	tggtcttcta	4080
cccagtgact	caaagcacta	aaagtccagca	taatcggaac	tgaagtcaat	agcatcgccc	4140
atttgccatt	cactgcagta	gcaaaaagtag	tactctgtgg	tgggttaatc	ggtttgaggc	4200
agtcctttaa	atgaacattt	gtgttttcatt	tttctgttat	tttcccgaac	atgaaaagac	4260
gataaaactg	aaatggaaaa	ggtaactgac	aaaagtgtgc	cttacctgtt	tccgccctga	4320
tttctgctga	ttcaagacta	ttctggctaa	actgattgga	ttctttttct	aactaggcag	4380
taggggatca	gaaatcacac	acggtaaccgg	ctgtgtttat	tctgagaggt	gctggggagc	4440
tttgggtctg	acttcttttt	acatgcctgt	cttctctttt	ggacagatct	attccagagg	4500
ggagottcac	cacttcattg	acggcttttaa	tgaagagaaa	agcaactgga	tgcgctatgt	4560
gaatccagca	cactctcccc	gggagcaaaa	cctggctgcg	tgtcagaacg	ggatgaacat	4620

ctactttctac	accattaagc	ccatccctgc	caaccaggaa	cttcttgtgt	ggtattgtcg	4680
ggactttgca	gaaaggcttc	actaccotta	tcccggagag	ctgacaaatga	tgaatctcag	4740
taagtggatt	acagaacaaa	aaaataaaaa	atgccagtaa	tgtcggttct	gcccccttga	4800
actaataaca	tgttgtttta	ttatacggct	ttgtcatgtg	ttggatgaag	taggtggctt	4860
aagctaggga	ctaggaagag	gaaaaacatt	ttttgagtc	ctattaacta	ttaggaaact	4920
tgatcaattta	aaagtatatata	tatatatgag	gagctacott	gagttttgaa	ttcaggatgt	4980
tacaggaaga	aatatatgtc	caatttctaat	ttatccaaaa	gcagttggga	gaattacagg	5040
gattgggtcca	gacatgctgc	gtatgcaagg	tatagccctc	atctgtggta	ctttggcagg	5100
gcttagactg	catcaaaaata	tttatagatg	tacattttgag	tgtacagtta	ggtactgatg	5160
tggaacattg	taagatcatt	gctagaaaaa	ctttgtcata	atttttcaat	attatttctaa	5220
gtgaataaac	gtaaagattt	tacatcttag	cttccttctc	tacagtaaaa	aaactatctg	5280
atctcttgat	cagtattata	gtagccacct	atcactttat	cttaacaaaat	tctcaattcc	5340
ttagggttat	gtgcttttac	ttcttttatt	tgattaaaaat	tgctgtcatg	acctctctct	5400
gcagagggct	gcatcatttt	ggtcattctc	aagtgtatctc	tttgagcaat	ttaagaattg	5460
ccataagatt	ctaacctctg	ctgtaaactat	ggttgtgtgt	tcttggttag	accactaaaat	5520
cttattagca	gttttaaaaa	ttatttcttt	tggtttagaa	gttaagacta	aatgctgaag	5580
tttttgtaac	ttttggtttt	gatatcattt	caaaccttaag	aaaacatttg	aagaaaaagg	5640
caaagaattt	ccacttaccc	tttaccagg	tttaccagtt	attgataagt	atatccattt	5700
gctttaccag	aaggctaact	tgtttttagtt	ctcattttca	cctttgagac	atttggaata	5760
aatatcaatg	ttaacataaaa	ttggaatttt	gacttttgatt	ttaggacca	tgaacaagcc	5820
aagtacttac	cctagtcata	tataatccaa	ctgtatgggt	atttggtatt	cattccacac	5880
ttcattttac	ttgatctccc	ttaagattgc	aagattgtgt	ttgcagtttt	tctgaaaatc	5940
tggggctata	aaagcatcag	gacctccccc	gtaggggagg	tctgtgtgtt	ggggtcctta	6000
cacaacaggt	tacccttgag	cttcaggaaa	agaactggct	ctcagttccc	cagttccagc	6060
ttaatgggtc	taattaggtc	ctgacccaaa	aggtggcagt	tcttttccct	catgtctctt	6120
cagcgctccc	cgagactctg	gagactctgt	catatcccta	gggctgagcc	tcccaggaa	6180
cattcggtctg	ttgtggcatc	tgtgtatgcc	atgcccagtg	ctgaggacct	agtaacaaa	6240
gacaaatgca	caggcacagt	ggcatttttg	tggaaactcg	attccagctg	tgcgtctcag	6300
aagaagcgca	cagctccctc	ctggctttct	taacatagtg	agccacttcc	acttaaggtt	6360
ctccttacat	tccttgagtt	taatcattca	tggattcaga	ggaaagtctt	ttgatttttg	6420
cttttcttta	aacagttcat	ttgaggtgac	ctaccccagt	gactttgcac	caaccaccaa	6480
gaaacttttt	tgcatgcttc	ccgcaccctg	tgccaatcaa	gggaagggtt	taaaggcctg	6540
gcgtttttat	tcctcaaaga	aaggtttttg	acagtatttt	aaggttcaag	tgcttctact	6600
ttgtgttcag	aagcaactgt	catatatact	gtgaaatgac	accttttatt	tatccctttt	6660
tatttatgca	gtatgtcccc	ttttattttg	gcagaatttt	ttctaaatgg	tggtttaaca	6720
ttttcaagca	catttcattg	tccaatatct	atagtaaaga	atgagagtta	acaataacca	6780
gtcacattaa	aacaagattc	ctgctgccag	ttgtgaaacc	ggttgtctta	ggcgtggcag	6840
ctgatgattg	agactgtgat	caggaataat	tccactattt	catcaggcct	aataggtaga	6900
ttgtgtctcc	aatgaactg	tgttgggttt	ccatgcttaa	agcacaatag	aggtgggtga	6960
agaatctcca	tgagggtcta	aatggcagtg	atggttcagg	cggtagagtt	tggagaagaa	7020
gggatttgaa	acaaacccaa	ggaaagaaaa	gtaagtagcc	agaaatcaca	aaatggcatt	7080
tttctaaaaa	caaaggaaaa	ggaataaaa	aactaataag	tttgaaaccc	ctacccctcc	7140
caaatttggc	agggggggag	gtattttttt	tctatctatc	taactaacc	atctagaaaa	7200
cagttgacca	aattatagac	ttctaaatgt	taatctgctt	tctcagtttc	agttgaaaag	7260
agactttgtt	ttgcctactg	cagaacttct	aggttctttc	ttatagtctt	ggggttctta	7320
ttatagatct	aaaaatgtgag	tcggcataat	taagccattc	ggagtcttca	gaagcagttc	7380
actcttgaaa	tgactccgtc	cgcctacagc	catttaagat	ttcagaacaa	aaacagatct	7440
tgattttctt	tttcatgtta	actcaagctg	ttgctgagtg	ggagagtcag	aaatgacacc	7500
agctccactg	attactcagc	tgctgaagga	tgatttttta	aaatgcacct	ttactgtata	7560
tggacttcct	aatttccacc	tgtagagcat	cttagggagg	ctaacatgtc	actctggatg	7620
ttcttttaga	ataagatgca	aatctatttt	tctgaaggca	ttagagatag	caaacattta	7680
ttgtgagttt	actatatact	aggcactgtg	ctaagtgttt	tgcatagaaa	gtttaaaaat	7740
ctggcttttt	tgttggccca	atcataagtt	tcatatcagt	tcaacattca	aattatatta	7800
aggtacttaa	gaagaatccc	tggctaaatg	tgaggggcag	tgccacagat	ggactgaaac	7860
tttatgctta	ttgcacattt	atgctattat	tatttgttga	attatagaac	caagggagtg	7920
tggaaagccac	tggaaaaaat	atgagactta	gatacataat	ttgagtaaaa	atgggtcaaa	7980
gtcatgaggg	taaagttttt	tgtattttcca	ttttatttca	gcggcatcgt	ttttaaaaat	8040
cattatgaat	ttgaccctat	atagatgttt	ccaaataatt	ctttttcacc	ttcataaaaat	8100

tccttctctgt	ggctgtgaga	tgccttgoot	atcagttttc	aagcttagtt	gtctttctca	8160
tcctttacca	tttttagcttt	aaaaaacaaa	agtgacaatt	agaacttctt	gcctgtctggg	8220
cctcactgaa	agaccgatat	tggcctgata	aggagatat	tattttgttt	tagtggtcttc	8280
agaaatccct	ctccctcagc	aagctttcca	tcacggcccc	cccgtcagca	tcttccctga	8340
tagcgttctt	ctctgtgttt	attctggggc	ttcaggctcg	cccaggagga	actgataacc	8400
gctggcagga	gataacattc	tctaaggggc	tctcaaattg	gaatcgaatc	cctcaagcca	8460
gtcagcctag	agaatacatt	taaaggggtc	agttctggag	tttcacagag	ttcatttcta	8520
gacctatcag	atagcaagtg	tggagtctct	tctcaactaa	attcaagcag	agacattttt	8580
tagacgatga	aggatatttg	cacaaaggct	tcagcatgat	ccccaaaacc	tgctgcctct	8640
gaaggcatct	ccacacattg	acagccaatg	ccttcagtg	gttccctagg	cagggtgtct	8700
ggcttgagtg	actgtcctcc	aataatcaga	gctcaaaacta	aacatcgtat	gttttacttt	8760
tggtttccag	gcaaggctga	gcagggaatt	ttcagttttc	cctgcccaga	tgggtgtttt	8820
ttcctgaagg	catcatttat	tgtgtagcga	ggagacaggg	ctggctgtgg	cagggatagt	8880
ctagaactgt	cctcattgct	gctgttccta	aatagtatct	ttaccaagta	ataacgtgcc	8940
gtctttggga	ataagtgtct	tcctcttagc	ctgttctgtt	ttcttgggtg	cgctaagtaa	9000
ttgaactggc	tcaggaagta	cctattgtgg	tttggcagag	gtgactgtca	cgcttgtgtg	9060
ctccaggggc	cagcactgct	gggatcctgg	ctagaccaga	cagagccttg	gtgaagtgc	9120
taggtctgtc	gcacatcgcg	aggaaggtgg	tattcacttc	gctaagctcc	ttggcatagg	9180
cagtttgaac	agggttttat	caaattcgta	ttcaacaaga	gtagaagcga	aaattgatga	9240
ctgtgtatta	cttgaaatga	gtcttaatct	ttcacattta	gttctcaggg	tatgctgatt	9300
tccttttaggt	aaaccatgaa	catcagaaag	actttttatta	acctatgaca	gggtccccc	9360
cccagttatt	ttccactcca	ttaaaatgga	agtttttttt	ttttttttct	tttttgagac	9420
agagtttttg	tcttgttgcc	cagtctggag	tgcaatggca	caatctcggc	tcaccacaac	9480
ctccacctcc	cagattcaag	cgattcttct	gcctcagcct	cccaagtagc	tgggattaca	9540
ggtgtgcgcc	accacgcccc	gctaattttg	tatttttagt	agagatgggg	tttctccatg	9600
ttggctcaggc	tgggtctcgaa	cttccgacct	caggtgatcc	gccacacctg	gcctcccaaa	9660
gtgctgggat	tacaggcaag	agccactgca	tcagcttag	gctatcttac	tcacgcctaa	9720
acagcaattt	tctatcataa	ggtctgtact	aatgaaaaca	gaatcaccca	aggctgtctg	9780
ttgttctgtc	tgtctgcca	ttgtccgcct	tgtctgagg	aggaaacgga	actgcacttt	9840
tgagtgtgtg	ggccagagcc	ttctagaatg	agagtgcgtt	ggaagccaga	tatgtggcga	9900
ttgtgtcgcc	agctgttact	caggttttct	caagaaggag	gagcaacttt	ggcagttttg	9960
cttcagttct	ctctagccct	ctgtgtaatc	gccccttttt	ctttatttca	gcacaaacac	10020
agagcagttc	aaagcaaccg	agcactgaga	aaaatgaact	ctgccc aaag	aatgtcccaa	10080
agagagagta	cagcgtgaaa	gaaatcctaa	aattggactc	caacccctcc	aaaggaaagg	10140
acctctaccg	ttctaacatt	tcacccctca	catcagaaaa	ggacctcgat	gactttagaa	10200
gacgtgggag	ccccgaaatg	cccttctacc	ctcgggtcgt	ttaccccatc	cgggcccctc	10260
tgccagaaga	ctttttgaaa	gcttccctgg	cctacgggat	cgagagaccc	acgtacatca	10320
ctcgctcccc	cattccatcc	tcacccactc	caagccctcc	tgcaagaagc	agccccgacc	10380
aaagcctcaa	gagctccagc	cctccacagca	gcccctggga	tacggtgtcc	cctgtggggc	10440
ccggctctca	agagcacccg	gactcctacg	cttacttgaa	cgcgctctac	ggcacggaag	10500
gttttgggtc	ctaccctggc	tacgcacccc	tggcccacct	cccgcagct	ttcatccct	10560
cgtacaacgc	tactaccccc	aagtctctct	tgccccctca	cggcatgaat	tgtaatggcc	10620
tgagcgctgt	gagcagcatg	aatggcatca	acaactttgg	cctcttcccg	aggctgtgcc	10680
ctgtctacag	caatctcctc	ggtgggggca	gocctgcccc	ccccatgctc	aacccccact	10740
ctctcccag	ctcgtgccc	tcagatggag	cccggagggt	gctccagccg	gagcatccca	10800
gggaggtgct	tgtcccggcg	ccccacagtg	ccttctcctt	taccggggcc	gcccgcagca	10860
tgaaggacaa	ggcctgtagc	cccacaagcg	ggtctccccc	ggcggaaca	gcccgcacgg	10920
cagaacatgt	ggtgcagccc	aaagctacct	cagcagcgat	ggcagccccc	agcagcgacg	10980
aagccatgaa	tctcattaaa	aacaaaagaa	acatgaccgg	ctacaagacc	cttccctacc	11040
cgctgaagaa	gcagaacggc	aagatcaagt	acgaatgcaa	cgtttgccgc	aagactttcg	11100
gccagctctc	caatctgaag	gtaggccttg	agagagagca	gtccaagggg	ctgtgagtgc	11160
atgcttctgt	ttgtatttag	cttgcttttc	atgggggtat	gattgcattt	gcagtagtat	11220
gagcccccg	ttggggatag	tgggtatgga	ttccgcctgg	cttttgccac	ttctagctct	11280
ttgactttgg	acaagtgact	tcccttctcc	tgattttctt	ctgaataata	aaaaaattag	11340
gggtttggac	tagaagatta	ggtgaaactc	cctgctagcc	tgtgattttt	gtgcttttaa	11400
gaaaaacacc	attctgaaaa	catgaagatt	tcttcttttt	aagactgtct	tgatgctttt	11460
cttaagatat	ttgcatcaac	acttgagtct	tggagcagaa	atgttaggtc	tcagagccag	11520
cttgagagca	gagctaacac	atgtggcttc	ttcccaggtc	cacctgagag	tgcacagtgg	11580

agaacggcct	ttcaaagtgc	agacttgcaa	caagggcctt	actcagctcg	cccacctgca	11640
gaaacactac	ctggtacaca	cgggagaaaa	gccacatgaa	tgccaggtgc	gcagtatttt	11700
ctgggtagac	cttctgacct	ttgtagaaaa	tgtctgtgag	tcacctcccc	atgtccctata	11760
tagccccgtag	ttaaagccaa	caccagattc	tgcgttgctc	catcctggac	tgatggcact	11820
atggtccttc	ccagtacttt	gtatctgctg	atgacttgag	atggcacagc	cagcttccag	11880
tgggtgggaa	aatggtaggg	gaaataaaca	gcccctcgtg	tgctgtgtgc	ccacatcccc	11940
ccgtttgctt	aataccacac	tggaggtgcc	acaaggaggc	ttctcacctc	ctaggttgct	12000
gggcgttggc	cggttaagcct	gcccctcccc	ttggcaactc	ttaatcttct	ggccttcctg	12060
tctcccttcc	ctgctgtctc	tctcccttac	actgtaggtc	tgccacaaga	gatttagcag	12120
caccagcaat	ctcaagaccc	acctgcgact	ccattctgga	gagaaaccat	accaatgcaa	12180
ggtgtgcccct	gccaaagtca	cccagtttgt	gcacctgaaa	ctgcacaagc	gtctgcacac	12240
ccgggagcgg	ccccacaagt	gctcccagtg	ccacaagaac	tacatccatc	tctgtagcct	12300
caaggttcac	ctgaaaggga	actgcgctgc	ggccccggcg	cctgggctgc	ccttgggaaga	12360
tctgaccgga	atcaatgaag	aaatcgagaa	gtttgacatc	agtgacaatg	ctgaccggct	12420
cgaggacgtg	gaggatgaca	tcagtgtgat	ctctgtagtg	gagaaggaaa	ttctggccgt	12480
ggtcagaaaa	gagaaagaag	aaactggcct	gaaagtgtct	ttgcaaagaa	acatggggaa	12540
tggactcctc	tcctcaggg	gcagccttta	tgagtcatca	gatctacccc	catggaagt	12600
gcccctccagc	aaccctactac	ctctggtacc	tgtaaaggtc	aaacaagaaa	cagttgaacc	12660
aatggatcct	taagattttc	agaaaacact	tattttgttt	cttaagttaa	gacttggtga	12720
gtcaggggtgc	ctgtaggaag	tggctgttac	ataatcccag	ctctgcaaag	ctctctcgac	12780
agcaaatggt	ttcccctcac	ctctggaatt	aaagaaggaa	ctccaaagt	actgaaatct	12840
cagggcatga	acaaggcaaa	ggccatatat	atatatatat	atatatctgt	atacatatta	12900
tatatactta	tttacacctg	tgtctatata	tttgcccctg	tgtattttga	atattttgtg	12960
ggacatgttt	gcatagcctt	cccattacta	agactattac	ctagtcataa	ttattttttc	13020
aatgataatc	cttcataaatt	tattatacaa	tttatcattc	agaaagcaat	aattaaaaaa	13080
gtttacaatg	actggaaaga	ttccttgtaa	tttgagtata	aatgtatttt	tgtcctgtgg	13140
ccattctttg	tagataattt	ctgcacatct	gtataagtac	ctaagattta	gttaaacaaa	13200
tatatgactt	cagtcaacct	ctctctctaa	taattggttg	aaaatgaggt	tgggtaatt	13260
gccaatggtg	gcagttgat	gtgttcattc	ctgggaccc	atcatttgaa	cagcattgta	13320
cataacttgg	gggtatgtgt	gcaggattac	ccaagaataa	cttaagtaga	agaaacaaga	13380
aagggaatct	tgtataattt	tgttgatagt	tcattgtttt	ccccagcca	caattttacc	13440
ggaagggtga	cagggaaggct	ttaccaacct	gtctctccct	ccaaaagagc	agaatccctc	13500
caccgcccctg	ccctccccac	cgagtccctg	ggccattcag	agcggccaca	tgacttttgc	13560
atccattgta	ttatcagaaa	atgtgaagaa	gaaaaaaatg	ccatgtttta	aaaccactgc	13620
gaaaattttcc	ccaaagcata	ggtggctttg	tgtgtgtgcg	atttgggggc	tttagcttgg	13680
gtggtgtttt	gttgttgggt	tttgttgctt	tttttttttt	ttttttttta	atgtcaaaat	13740
tgacaaaaca	tgggtgctcta	ccaggaagga	ttcgaggtag	ataggctcag	gccacacttt	13800
aaaaacaaac	acacaaacaa	caaaaaacgg	gtattctagt	catcttgggg	taaaagcggg	13860
taatgaacat	tcctatcccc	aacacatcaa	ttgtattttt	tctgtaaaaac	tcagattttc	13920
ctcagtatatt	gtgtttttac	attttatggg	taattttaatg	gaagatgaaa	gggcattgca	13980
aagttgttca	acaacagtta	cctcattgag	tgtgtccagt	agtgcaggaa	atgatgtcct	14040
atctaattgat	ttgcttctct	agaggagaaa	ccgagtaaat	gtgctccagc	aagatagact	14100
ttgtgttatt	ctatctttta	ttctgctaag	cccaaagatt	acatgttggg	gttcaaagt	14160
tagcaaaaaa	tgatgtatat	ttataaatct	atttatacca	ctatatcata	tgtatatata	14220
tttataacca	cttaaatgtg	gagccaagcc	atgtaaaaaga	tctacttttt	ctaagggcaa	14280
aaaaaaaaaa	aaaaaaaaaa	gaacactcct	ttctgtgact	ttgcttaata	cttggtgacc	14340
tcacaatcac	gtcggtatga	ttgggcaccc	ttgcctactg	taagagaccc	taaaaccttg	14400
gtgcagtggg	ggggaccaca	aaacaaccag	ggaggaagag	atacatcatt	ttttagtatt	14460
aaggaccatc	taagacagct	ctattttttt	tttgccactt	tatgattatg	tggtcacacc	14520
caagtcacag	aaataaaaaa	ctgactttac	cgctgcaatt	tttctgtttt	cctccttact	14580
aaatactgat	acattactcc	aatctatttt	ataattatat	ttgacatttt	gttcacatca	14640
actaatgttc	acctgtagaa	gagaacaaat	ttcgaataat	ccagggaac	ctcagagcct	14700
tactggtcct	ctgtaacttc	caagactgac	agctttttat	gtatcagtg	ttgataaaca	14760
cagtccttaa	ctgaaggtaa	accaaagcat	cacgttgaca	ttagaccaa	tacttttgat	14820
tcccaactac	tcgtttgttc	tttttctcct	tttgtgcttt	cccatagtga	gaatttttat	14880
aaagacttct	tgtcttctct	accatccatc	cttctctttt	ctgcctctta	catgtgaatg	14940
ttgagcccac	aatcaacagt	ggtttttatt	tttctcttac	tcaaagttaa	aactgaccaa	15000

<210> 66
 <211> 46340
 <212> ADN
 <213> Homo sapiens

<400> 66
 tatttttacttt cagtaacaga aaatgaaaga aatgtttttaa tgttgctgat tgtattacct 60
 tcaggatcaa tagcagaagg acaaaacttct ttgaggagat ctccctagtg gtgcaactgt 120
 ccactctgcag ccacaggacg aaacagcttc tgaatgaaa gtcctttcagt cgttgtctat 180
 ttgaaaaagg aaaaaatgat tcaagcaatt aagtcctttgt tgctgccaat tacaaaattta 240
 tataatcataa acttttatgtt ggcatttaggt gcctttttagt acgggtgttag cataattaca 300
 caacatcaca gatgtggtat cactgtgaaa aatgtttaac atgataaatt caggtaaattc 360
 taattctgag gaaacagaca aatccaaagt tgggtgggac attcctaaaaga taattggctg 420
 ggacccttca aaaactttaaa gacattaaaa agcaaaacac acaaaaagat atcaacaaaa 480
 gcatttttttc tcagtatctc ttaaagagac taacaaagca aatacaaaac ataaaccatg 540
 gctgaataact aaattgaaga aggacatttt tttagaaatcc aactatgaaa cacagttttg 600
 ggataaatgg ggaaatacag aatggacaac tgataatatt attgagttaa tgtcaaattt 660
 cttaggtaca ataaggacaa tccttatttt taagaaattc attgttcaag tgtttaggaa 720
 agaagtgcc aatgatatccaa aacttaattct tctttctctt tttttggaga cagagtctcg 780
 ctctgccacc ccggctggag tgcagtggcg cgatctcagc tcactgcaac ctctactttc 840
 caggttcaag tgattctcat ggctcagcct cccaagtagc tgggactaca ggagtgcgc 900
 accatgtcca gctaactttt tgtattttta ctagagatgg ggtttcacca tgttgcccag 960
 gctggtctca aactcctgag ctgaggcaat ctgcccgtt cggcctccca gagtgttagg 1020
 gttacaggcg tgagccaacc gctcctggcc ccaaaactta accatctaatt ggttgagaga 1080
 gagacagaga gagagagaaa gagagagaca gagaatgtgt gtgtgtgtga agacaaagca 1140
 aaaataaaaa aatattaact aatggtgatt ctaggtagag ggtgtatgat ttttagtagt 1200
 tcattatttc aacttttcga taggtttcac aatttccaaa acagcagatc cagccatttc 1260
 atctgacaaa aactgttagc agcactacat cgtaatttat tgctaataat ctctattgtt 1320
 tactcttaaa attgtttcat ttactaaatt tccttagtga tgatggaggc tttatcatga 1380
 cagagtacag aggctctgaa atgagccagt gtctatgaag agcaccactg tttgcaagat 1440
 ctatgatctt gtacccagtt tccttttatct gttaatttgg gacattccat atctcttgag 1500
 tttgttgtgg aaataaatga gcaactttgc caaccacaga gtaaataaat aaatgttaaa 1560
 gagaataaaa gcatttttac ctctctctc cctcttaacg gttatttcac ttttaagatg 1620
 taaattttta gctttctgag atgaaaaatc attaaaactt aacaagaaca gagaaatgcc 1680
 atacatacat attttttgtt tgcttgtttc ctgagacaag gtttcaactc gtcaccagg 1740
 ttgaattgca gtggtgcaac cccaagttg caatcctcca cctaagcctc cagagtagct 1800
 gggactacag gtgtgagcca ccatgctcag ctaatttttt tacttttttg tagaagggg 1860
 tctcactatg ttgcccaggc tgccctcat tttataagaa tatgacttca aacacttagg 1920
 cattagcgac aaggttttgt ttttgccttt taatgacaga ggtataacct aacatatttg 1980
 acacaactgt tagagatttg gtttaaaaag aaatgacat ggatgaagct ggaaactatc 2040
 attctcagca aactaacaca ggaacagaaa accaaacacc tcatgttctc actcacaact 2100
 gggagctgaa caacgagaac acatggacac aggcagggga acatcacaca ccaaggcctg 2160
 tcggggagta gggggctagg ggagggatag cattaggaga aatacctaac gtagatgagg 2220
 ggctgatggg tgcagcaaac caccatggca catgcatatc tatgtaacaa acctgcacat 2280
 tctgcacatg tattccagaa cttaaagtat aatacaaaat gaaaaaataa ataaaaataa 2340
 gtagaaaaaa taaacatgta agcatgtgag ctgcctttcc taattctatg tttatgtatt 2400
 cactgaatac atagtatttt aaaatagtaa tccaataata tatttgagtg tttgtgacaa 2460
 gtatgaaaat tgtaattttt aaaaaatctt gataatatgc attgaatatg atttaattca 2520
 cttcactatt tgaactcttt agggattatt tttaaaaata tgattgatat cctttgatat 2580
 gttttggctc tgtgtttcca tccaaatctc atctcaaatt gtaatcccca ccgctctagg 2640
 gagggactgt aatccccatg tgtcgaggga gggggtgat tgggtcatag ggggtgtttt 2700
 cctcatgttg ttctcgtgat actgagtga tttctcatgag atctgatggt tttaaaagtg 2760
 gcagtttttc ctgcactctc atctctcttt cctgctggct tgtgaagggt cctgcttccc 2820
 tttctgccat gatttttaagt ttctgaggc cccacaaagc catagcgaac tgtgagtcaa 2880
 ttaaaccctt tgcctttata aattatccag tctcagatat ttcttttaag cagagtga 2940
 acagactaat acattcttca atttaaaaag ccatactttc tcatacaagt tgaaccaag 3000
 aacaatatca tgcataatca agtgattaac tgtgtaaaga taataagggt gaggagtcca 3060

gagaagaaaa	gaaatgaata	gggaactgta	gtgataat	aaaatagcca	tccctcactc	3120
aggggtttttg	atcttcaggc	catgaagaag	cttttaatgc	tttttagcaa	aggaagtaat	3180
gttggtgaaa	ggctttttct	gacgactaat	ggaaagcagt	gctatgtatg	gtgacttggt	3240
tatgaaccaa	aaccagaatg	actggtgaga	ggctgactga	atacagcaag	cttatgtgaa	3300
gacaactgga	gctggtgcag	tggaaaagga	agacagcagg	actgtaccca	caactcaaag	3360
aaaaaagtc	gaaggtacct	cccgcagtcc	aacctgaaaa	caacaaagtc	aaaggaatct	3420
tttcaagaat	ttggagctct	cattcatatc	ctaattagtg	tatgaaatgt	gaggtggctt	3480
tgtctataatg	aaattacctg	gaatattttct	aacacaaaga	aataataaat	gcttgagggtg	3540
gtgaatatcc	tcatttgatc	attacacatt	gcatgcttat	agcaaaagat	tacatgtacc	3600
ccataaataa	ttgcaactat	tatgtatcca	taataattaa	aactaaaaga	ttaaaaaatta	3660
cctgaaaaaa	aatgctaacc	aggaaaaggcc	aactagtctt	ggttacatac	taaaaaacag	3720
aaattcttct	ctaaccctcac	tattggagaa	atatacctgtt	atcttttttat	atctttttttt	3780
tcaccctttc	ccaaatctga	gcaagtatta	taaagggtata	accttcaaca	atctttttatg	3840
atgaggtatt	tgcttactgg	ggacaaagcc	ccagtgtctat	tacatagtgt	agctaaacgc	3900
tgtagaatgg	taaaaacaag	aaaatgctca	gcaaagtgtt	gtttctcatt	taatgaaaat	3960
cttatttttaa	aacacaaaaa	ctcaatatac	cccaaccaaa	aatctgatga	acattttctg	4020
tttaatat	attatacagt	accttttaaaa	acgtaatat	cttattctta	aaaattttagt	4080
gtgctagcaa	atagcaatta	agtacctaag	tcaatcagga	cgacaaaaaa	atactcaatt	4140
tggggagtta	gttacttcta	tcactctgaat	gcgtccctcc	aaaattcatg	ctgaaacctta	4200
ttcctcatca	tggcagtatt	aagaggtgaa	gccttttgaga	ggtaattagg	tcatgagggc	4260
agagtcctca	agaatgggat	caatgctctt	ataaaagagg	cccagggag	cttgtaaggc	4320
ttttgccctt	tctgccatgt	tgggggggtg	gggggtggggg	cgcagcaacc	agtgtact	4380
ctgaagcaga	gagcagccct	caccagaaac	cgaatctgtt	gaagccttga	tctctgactt	4440
cccagcctcc	agaactgtga	gaaataat	tctgtttgtt	ataaattacc	cagcttaggc	4500
tgggcgtggt	ggatcacctg	aggtcaggag	ttcaagacca	gcctggccaa	tatggtgaaa	4560
ccccatctct	actaaaaata	cagaaaatta	gctgggcata	gttgtgggag	cctgtaatcc	4620
cagctactca	ggaggctgag	gcaggagaat	cacttgaacc	cagaaggcag	aggttgtagt	4680
gaatcaagat	catgccattg	aactccagcc	tgggcaacaa	gagggaaact	gtctcaaaaa	4740
aaaaaaaaa	aagtacacac	tctaacatat	tttggtatag	cagcccaaat	ggaatggact	4800
aagacaatta	cccttaaaat	aaaagctccc	atagagagat	catgcattca	agacagagg	4860
ttcttaagg	caatgggaat	ggaggacata	ttcctgcaaa	cttttcaaca	gctctcatta	4920
gcccgatgtt	agagctctgc	aaagaagact	aaattatact	gagaaatatt	tttaaatctc	4980
cacaaatagg	aatgctgtaa	acgttgattt	agtatatata	aaattagaca	agactaacaa	5040
tatccaatgc	aatctaaatc	ttaggttgac	agacaagaaa	gccactgcaa	acaggaatat	5100
accacaatac	ctgatcttgc	cacatatttg	taaatatgca	aagtatttca	ataactttca	5160
agaaacagta	ttactctcat	gagaaataac	atgatgtaa	tcacctttga	aactgtcctt	5220
gttacttttt	caaagtgtatg	ttagtcattt	cttaacacca	aatgaaatga	aaaactgagg	5280
tggtaatggc	tggctgctcc	catctctcct	ctactcatgt	gccttcacca	atacagcaat	5340
cattttttct	tatatgggaa	atttacagtg	ttgatatagc	tcagagatat	attgaagaaa	5400
agcagaaaaa	cgaaacttat	aaacattttta	ggaaacctta	tgtattttct	taaatagttc	5460
aagtgtaaaa	cttagaattc	ttataaataa	tgtgtgttac	agctatat	taaatgggtg	5520
ctcatgcctt	taatcccagc	acttcaggag	accgaggtgg	gaggagagct	tgagcccatg	5580
agtttgagac	tcaccggggc	aacacagaga	gacctcatct	cttaaaaaaa	aaagaaagaa	5640
agaaagaaat	gaaatgcaa	gaaaaagtct	ctattttcaa	tgtagccagt	agagccaata	5700
ggttaaccaa	tattaacatt	aacgttgata	aaacaagaaa	tgatgattta	ctataagctg	5760
aaaatcagac	aatgtatgga	ctttaagagt	aacaggcacg	atcatcacia	acttaaatca	5820
ggtttgagtc	ctatgagtta	tatacagtta	catgatgcaa	caaaagatgc	cagccagttg	5880
ttaaagagta	ttagattcgg	ctgggggtgg	tggctcatgc	ctgtaattcc	agcacttttg	5940
gaggccgagg	agggaggatc	acgaggtcgg	gagtcagaga	ccagcctggc	caatatagt	6000
aaacctgatc	tctactaaaa	atacaaaac	tagtcaggca	tgggtggcacg	tgccgtgta	6060
cccagctact	cgggaggctg	aggcaggaga	attgcttgaa	cccagggggc	ggaggttgca	6120
gtgagccgaa	atcgcgccac	tgactctag	cctgggcaac	agagcaagac	tctgtctcaa	6180
aaaagagtat	tagattcaag	tccgtgtttct	gtcattttatt	atggaacctat	ggacacacac	6240
acctatcttt	cctgaacctc	agtttttttca	actgcaaaac	aggaatatata	acatatgtgt	6300
atatatacat	ctgtgtaaac	acatatgtgt	atatatacat	ctgtgtaaac	acatatgtat	6360
atgtataaat	ggagataata	cctacattat	agtttctgag	ataataaaat	gcacaacaca	6420
attctgacac	ataacaattt	gtaacttaaa	acataccatc	accagggcca	ctagtttttag	6480
aacactgtaa	tgcatagtct	aatttaatac	tatgcaaaact	gtgttcactc	aaggtttttat	6540

ttccttttaa	tttcattcat	ttactcttca	gttgttttga	agctaaaaag	tccagaatca	6600
tgaaattcag	aagtttacgt	tttaattgtt	ttctatatgg	caaggaaaaa	aaaaagggca	6660
aagtcatttt	aacactactt	tcaaaatcag	cctagaactt	aacactaaag	gcatgaccca	6720
taaaagggaa	tactaataaa	tagacttaat	taaaattaaa	caacaacaac	aacagctaag	6730
cttttgttct	gcaaaagatc	ctgtgaagag	aatgaaaaca	taagccgcag	gctgggagaa	6840
aatatttgca	aaccatattt	ccgagaaagg	tcttgtgtct	ataatatata	agaactccca	6900
aaattcaaca	gttttttaaa	aaagcaaata	atccaattag	aaaatgggca	aaagacatga	6960
acagacattt	taccaaagag	aatatatagg	tggcaaataa	gcatatgaaa	acatatctca	7020
cacatcatta	gccattaaag	aaatgcaaat	taaaaccaca	atgtgatata	attacacacc	7080
tacaaaata	tccaaaataa	aaatttagtg	taacaccaa	tgctggtgag	catgtggaaa	7140
aatagtcctt	cacacactga	tggtacaaat	gcaaaacagt	acagtccctc	aggaaaggag	7200
tatggcagtt	tcttacaaaa	ctaaacatgc	acttaccata	tgaccaagta	attatactct	7260
tgaatattcc	cagaagttaa	aatgtcttct	ccaaaaaact	tatacatgaa	cgttcatagc	7320
tgttttatct	gtgagagtca	aaaacagaaa	gcaatcccag	ggctacccat	taaaacaggt	7380
gaatgcttat	aaactgactg	taataggtct	gtcccacgga	atactactca	gcaataaaaa	7440
ggaacaaact	actggtatat	gcaacaactt	ggatagatct	caagggagtt	atggttatgtg	7500
aaaaaagtca	atctcaaaaag	gttacacact	gcatgactcc	actgatataa	cttagtgtaa	7560
atgacaaaaa	ttttagaaat	ggaaaacaaa	ttagtagttg	tcagaggtta	gggaagaaat	7620
gcagtaagggt	aggtggctgt	ggctataaaa	gggtagccta	agagatcctt	ctggtgaaac	7680
gggtatatatt	tgaatatagg	gtgaattttac	atatgtgata	aagattgcat	agaactaaat	7740
acacacacac	agtatatgta	aaactaagga	aatctgagta	aggtttgtgg	attatatttaa	7800
tacaatttcc	tggttgatgt	actgtactgt	aattatgcaa	gatggttagaa	ttgggggaaa	7860
ctagatgaag	ggtagttaga	tctttctgta	ttatttctta	caattgcatg	tgaatctgta	7920
attatctcaa	aataaaaaatt	tttttcaaaa	tttcaaaaca	actagtctag	agctttgtta	7980
atcaaagttt	tctctgagga	cctgtagcat	tttggttatc	acctggatct	tattaaaatg	8040
tagattctca	ggctgcatat	tggaattcct	gaattggaat	ccgcatttta	acaagatttc	8100
caagtgattc	atgttttaaag	tttgagaagc	actagtctac	aacaatgact	tttaaccttt	8160
caacctactc	taacacactt	gaaggccata	acaaaattca	catcaataac	agttgtctcg	8220
ttggacagtg	actctcaaca	caaatgagtg	aggaaagggtg	gggactcaag	actcaggtag	8280
caggaaaagc	cccttaggtg	atcctgatga	aatgttttct	ccatcctggc	tgaaaaaccc	8340
agaacagtca	attaaggctc	aaaacaaaag	taatgtttat	aatactggag	atctttaaaa	8400
ggcagataat	atatactata	acagagcaaa	ggtaattatt	acaatgtata	aatcttataa	8460
gaacccaaat	cagaattaaa	atcactaagc	acataatgaa	aatcctttta	aaagtataaa	8520
aatgaatgta	gtctaagtaa	atactaataa	tggcagttat	agtgaagaaa	gctctagagt	8580
cttttactct	tcataacttc	tagtcacaaa	catctatttc	caaaactgac	ccttcgtatt	8640
tcaaataatt	tatggcctgg	tacagtaata	agagcatgat	atttaaagcc	agtcagaaga	8700
cacatattct	agctctggat	ggcacttgat	gacgatggat	tcagcttatg	gttccaatcc	8760
cagctctgtc	aattagtagc	tatatgaccc	tagtcaataa	cttaaacctt	cttgtgttac	8820
tttgtgtgta	attgtatcat	ctataaaatg	aggatatata	cagtatatat	ctcatagatt	8880
tttttgtgaa	ggttatacaa	ttaattcata	taaatgtatt	agaacaatgt	ctagcacagt	8940
gaattctcaa	tgagtgttat	aattgttctt	tttaaatgtg	acttgactct	caacagaact	9000
ctactgaatt	ctaatatgta	ttctgtattg	agctgtcaaa	aaaaataagg	attataataa	9060
catatactat	tcttgtagtc	aaccctgtta	ctatgttatt	actagtgtca	gttttgttgt	9120
tttggtcata	catattgttt	tacatacatt	aagaattatt	agaaatgttg	gtttatttaa	9180
aatgaccatt	tatggctaga	agggtatata	tctggctcac	tgactgtgga	gtcaatgtcc	9240
ataaagagga	ggaagaatgc	catcagagta	aaaggagatt	ctattcactg	aaacaaagt	9300
ataaaaagct	atgaaagaga	aaaacataaa	ataaccaaag	gggtgaaact	taacagatgc	9360
ccagtagatg	cacaatgcac	tgggttgtaa	aacttaaaat	ggccttaatt	aaaagccaag	9420
cacggatgga	ggtgctgggg	gagtctccta	cggacacagc	aggcagaatg	taacaatgac	9480
aaggggctca	agtttatatta	aaaagagatt	ggacaggccg	ggcgtggtgg	ctcacgcctg	9540
taatcccagc	actttgggag	gctgaggcgg	gtggatcatg	aggtcgaggag	ttcgaggcca	9600
gcctggccaa	catggcgaaa	cctcatctct	actaaaaata	aaaaaaatta	gccgggagtg	9660
gtggcgtgca	tctgtagtc	cagctactca	ggaggctgag	gcaggagaat	cacttgaacc	9720
tgggaggcaa	aggttgcagt	gagctgagat	catgtcactg	cactccagcc	tgggcaacag	9780
agtgaactg	ctcaggatct	cccaaagacc	caaatccctg	taaactgaat	gcataatatc	9840
atgtgtcca	gtgaggotta	gatggacatt	ctagtcttct	tggttgagct	gaagaaacaa	9900
atattatatt	gataatttat	gtatgttgta	tttttcaagg	tatagcaaca	agtttttatt	9960
catcagctac	tttgtgtgtg	tgctttgttt	ttaagtcctt	tgaaacagga	tggtgattta	10020

ctacattttat	aagtaaaatt	tattttgattt	acaaggggttg	cttaagtgtga	tcacaggatt	10080
tcacttggtta	tattttgcagg	tgcttaaaaa	atcagctata	ctaaactata	actggaatta	10140
gcaaagttca	tttattgatt	aatcaagaat	ataattagat	ttgcctaact	atataagtag	10200
tactatgtgt	tattttaagaa	ttaaatctag	aaaagggatg	gactctggaa	atatcaagaa	10260
gtgaaaaaga	ctgctctcat	ttttgtacaa	caattactaa	atttctaagt	agcatttaatt	10320
gaactgaaaa	ggcatttttag	aaaaactaga	ttttacaatt	tataactcta	ataaaacaca	10380
actaactatg	agtgtgcttg	ttcatgcccc	aaagctacct	tccaaaatta	aaaaccctat	10440
tggatggctg	ggtgcagagg	ctcatgcctg	taattccaagc	actttgggag	gccaaaggcgg	10500
gcggtacc	tgaggtcagg	agttcgagat	cagcctggcc	aatatggtga	acccgtctct	10560
aacaaaaata	caaaaaattag	ccgggcgctg	tggcgggtgc	ttgtaatccc	agctactcgg	10620
gaggctgagg	caggagaatc	acttgatcct	gtgagcgga	gggtgcagtg	agctgacacc	10680
gtcccactgc	actccagcct	gggcgagagc	ccagagcgag	actccgtata	ttaaacaaaa	10740
caaaacaaaa	ctcaaaaaac	cctattggca	attactaggg	ccatcaaatc	agtatatatt	10800
cacttgacac	acaatttttga	gataatgaac	cgaaactact	atttttgaaa	atattacata	10860
ataaatatta	gtgaagcttc	attgctgaaa	tggtgacaaa	gatgaatagc	aataaaaactt	10920
ttcttataga	tcttttagcaa	aaacaaaaaa	accccaagca	tactatggta	cattacttta	10980
gagaatcaag	tagctgctag	ttgagtaata	gtgtaataag	gcactacaat	gatataaaca	11040
aatatacaaca	aagaatattg	tttttatttc	ctgtccatgt	tttaaaaaag	ctttggtttt	11100
acctatgttt	aacaaaagca	taggtacaac	aacgactact	actactaaca	tataagtagc	11160
ctggatagaa	ttatcttaat	agtagtacct	aaagtgcagga	tctctaagta	atgatcagaa	11220
ggcaggaata	aatttttatca	gaaatcttca	ttcattacat	atttactatg	cattttaccag	11280
ggtatcacta	tgctaattgga	tacaaagata	aataacatgc	aaacaactgt	aatacagtg	11340
tatgtgataa	cagaaatatg	tacaaagcac	tatgaaaaaa	attacaaaagc	ttgagcacaa	11400
attttaactc	tggacttact	ggcattttaga	gcaaaaccaa	aacaatccta	actggttaatt	11460
ttcattttct	aagagttgga	agctatatca	gtaggtacaa	agtaaaatat	gctaattgtg	11520
gtagaaagta	aaatattaca	acagtagaga	atttcaaaaag	aagataaaaa	taatggaggg	11580
aatatagaag	gtcttcaagc	ttccagcttg	aaatacatat	tttttttttaa	atagagaaaag	11640
agataaagtc	atttgagtat	tcagagggca	gactgaatat	aatggtactt	ctgagaaatt	11700
agtgataaag	gagagaaaag	tggactaaaag	gcatatgcat	atagagcttg	gaatgtcaaa	11760
tgtagtggaa	ataacaaaag	tttggttgga	atcccaactc	ccaacaacgt	actgtgtatc	11820
tagagcaaat	tacatcaacc	tttgggagta	ctgtttctga	atctgaaaaa	tgaggaaaac	11880
ttatctttga	acaattgatg	tgataattaa	atgagatata	tgaaatatct	aatgtaacaa	11940
gtgcttaaca	atgactagtt	cttttcattc	ctctcttgaa	ccattgtgaa	acgtagaacc	12000
aagaaaagta	acagtattta	gttggtacag	aaccatttaa	gagagaataa	aaaataactg	12060
gtattctaac	ttcagtttcc	tttgaagtct	tgtaattgag	aataaatatt	atgtggcaca	12120
aagaaaaaga	aaacaggggt	ttacacagga	tatgctgcca	gactttacca	acaatgacac	12180
atgatatctg	cttcaactgt	cccatgcata	tttggcttaa	gatataattca	tgcatatcaa	12240
attttacatc	acatggtttt	caaaagaaga	ttcattaaaa	ttagcttaag	aatgtacaca	12300
atatacaata	cctcattaaa	taaaaaagaac	agaccatttc	caaatgaatg	cttttagagc	12360
tttacagtaa	acagtccttt	ggtggtagaa	agagggggaa	cagagagggg	agtgggtggg	12420
agtctgtagc	acttatcaga	ctacttttat	cctttatgta	gagaaatagg	agagttgaaa	12480
ataagcactt	tctgtactta	tgttgagagt	ctgaagccca	cttttaatat	tcttgacaac	12540
actaaaaaat	aataattaac	atttgaaaag	ctgtcattat	tatagtcagg	gacacttaat	12600
ctccaaaagga	gaagtttctt	aattgatact	atgattaaat	aaaagcatcc	atcagaatta	12660
tatccacaat	ctggtttgga	gtttatgttt	tgtcttattt	aaattgttat	acttattata	12720
attctgtcta	gacagtgcc	aatgtacttt	gtcatacaaa	cacttgaggc	aaattttctt	12780
caaataagcg	caacactttg	tttcctcttc	gtatcctttg	actgaataac	gtgtggtaca	12840
gagaagtaat	acttcccttt	cttgggatcg	agatcaattt	gatgcttggt	ataagcccat	12900
ttacagaaca	aatgggtattg	cttttaaat	tttatatgaa	cttatcagta	gactagccaa	12960
aaaagaagct	tcatataaaa	gtgctaggat	tgatattctt	agtaataatt	aggtaaattc	13020
tctaaaaatt	tctcccaaaa	gatctgaaaa	atcataccaa	gggaagtata	gttttaaat	13080
cattatatat	aatagcttta	aaatatcttt	gctaattcta	cccaaagcca	cactaaaaag	13140
actaatacaa	aaagaatgta	atttaataaac	tattttcttc	tgaagaatca	aagggcactt	13200
ctgcatatga	acatgtttta	tccttttggt	gtacttacat	aaaataatta	agaaacactt	13260
ttaattagta	taaacaaaga	aatcaaaaata	gcaagaagaa	atgtctgagt	aaaagcagct	13320
gtgctgacct	caaaagtga	attctgttct	cttgatgcc	agttaagtgt	ctaaccag	13380
gaaaagtgat	tctaaacctg	ggctaggagc	tagtggagct	cttcaaacag	tctcacctac	13440
cctcaccct	caaggaatgg	tctatgggtt	ctgtggtgaa	cgctaaagtt	tataacatgg	13500

gaatattttat	tatttttgttt	ctaacacaaa	taattttttaa	aaattttatto	tactaaaagta	13560
acatcaaagg	gaaattttcat	aaaaatttctt	ttgaaattttt	tagaagtagc	aaataaaggc	13620
aagtgataaa	tatttttacag	atttcaccac	ttacgtaatc	tgatcaacaa	atttttaaaaa	13680
catagcactt	gaatactatt	aaaaatatat	taaaaaggta	acatagtaaa	actataaaaa	13740
tctttaaaaa	aaatataaga	ggaaaccttc	gtgaccttgg	attaggaaat	ggtttcttao	13800
atacggcaac	ctaaaaatac	aagcaaccac	agaaaaaaac	agacaaaactg	gacttcatca	13860
aagttaaaaa	cttttgttct	tcaaatgaca	tcatacaagaa	aataaaatccc	acagaatggg	13920
acaaaatatt	tgcaaacccat	atctgataag	agaccactat	tcagaatatg	taaagaattt	13980
gtaaaactta	taaataaaaa	gttaaagaag	tcaatttttaa	aatgagcaaa	ggatctgaag	14040
acaatttctc	taagaaatac	gaatggctag	ttaaattgcat	gaaaagatgt	ttagcatcac	14100
tggtcattag	gaaagagcaa	aaaccaaatt	gatataactcc	ttcataccca	ctaagactgc	14160
tgtaattaaa	actatagaaa	ataagcggtg	gcaaggatgt	ggacaaaatg	gaacctctcc	14220
catacactga	tggtagaaat	gtaaaatggt	gcagatgctt	tggaaaacag	tctgacaata	14280
ccccaaagg	ttaaactgtg	aattaccatg	caaccagca	attctactcc	taagtatcta	14340
cccaagagaa	atgaaaatat	atgttcacca	aaacatttgt	acataaatac	taactgcagc	14400
ttttattcat	aatagccaaa	aagtggagac	aatccacatg	tctatcaatt	ggtgaattga	14460
taacaaaaat	gtgggtatctt	catacaacta	ttactgggcc	ataaaaaagaa	tgatgtattg	14520
atacatgcta	caaaatgaat	gaaccttaaa	aacaatatgc	aagcaaaaaga	aaccagacac	14580
aaaaggccat	atattacatg	atgctaatta	cataaaatgt	ccagaaggga	gaaataaatt	14640
agtagttgcc	aagggctgga	gggaggggga	atgatataag	tgactgccaa	tgggcatggg	14700
gtttcttttt	agggatgatga	aaatgttctg	aaattttatc	acgggaatgg	ttgcacaact	14760
ctgtgtaact	tagaattcag	tgactcctaa	aaccaatgaa	tagcatgctt	taaaagggtga	14820
cctttgctga	gcatagtggc	tatagtccta	gctacttggg	aagctgaggc	aagaggatca	14880
cttgagccag	gagttccagg	ctgtactgca	ctatgatcat	acctgtaaat	agccaccata	14940
cacaccagcc	tgggcaacac	agaccatgtc	tctaaataaa	taaacaataa	aataaataaa	15000
agggtagacct	ctgtagtatt	gagattatac	ttcaagtaag	ctgttattaa	aaaaaaaaaa	15060
gttatcatat	gggtggcagg	ggaaatcatt	ctgggatgat	ggctaacttc	atcagtattt	15120
gatttatacc	tatgcatcat	accttatggt	tgttttatgc	attttgtggg	ttttttaaaa	15180
aaatttatatt	tcataaaaaac	aaatttttaa	aaaatttaaag	tcaagaacc	caaaacaaca	15240
aagatcagag	atacattttct	accttatcaa	ttcagaaaaa	ttacaagtgt	ttttcttaaa	15300
aattgtatag	catcatgggtg	attttaagtt	acctgtagga	atttaaataa	ctttgtctta	15360
actgttcacc	aaaactcatt	taatattcat	gtttctgatac	tgaaaatgaa	gctgaaaagt	15420
tttgaaatta	caatatgcta	gtttaaaaag	gtttactaaa	atacataatt	tcattataag	15480
gagtaatatg	aaataaaaagt	atcaaatatg	ggacacattaa	aaatgtcctt	aaatacaaat	15540
tgctaccac	attgtggact	cactgcgtcc	actgtttgcg	agcttttcca	gaacgctcgc	15600
caccagttag	ggtagccaag	aactcctcat	cttcactttc	ttcctcacta	gcttgggaacc	15660
tctggattcc	caccacact	gctgtgacct	gaatggggaa	gagaaacgcc	atagtaaggg	15720
aactcttcct	tttatagatt	tctgaattag	aatctggcat	tacaaaagaa	caatgttata	15780
aatccaggtc	agagtttata	gttctatttc	actattactt	atatggcttg	tccttaggaac	15840
ttaactatta	tttacaatgt	aagtacattt	ttccacaaaa	aaattcaaaa	ttttggaata	15900
caatatctga	agagagaatg	gtctattgaa	tccaaaagtag	gctgatacat	cccaacagta	15960
tttcagattg	agataataat	aataaccacca	attcatcaag	tcaaattata	tgcttatttt	16020
ccacaatgga	agtttttaaa	tagtataaac	attttaatat	atagcaggct	taacttatga	16080
ttattaaaca	gggttctaag	aaaatagtat	acatcaataa	ttaatgtgct	tcttgtataa	16140
tttaggtgac	aatttatcca	tctgagaaat	gcaaaagaga	cttttggtag	gggttgagta	16200
aggagcattc	tgtgtcaaag	aattcactag	caaaagaggg	tatactgtag	ttacaagcta	16260
taatcactgt	acttattttta	aatccctctt	cagaaccagg	tcttaaaaaga	tgataaacat	16320
ggcctcatga	ataactatca	accaaactat	agaaaagagt	gcaagagtgt	ggtgttctaa	16380
cttaaaaatat	ggtgttttat	tcaaataatt	ttattttaagg	ctccaaaagc	agcagcctca	16440
ttccccagaa	atcatagtta	aatgaaatct	tccttactaa	aggaaaaatg	aatcacataa	16500
tttaacgtga	acatttttaa	aacactctaa	agcaacaaaa	ctattcaatt	gtatgtgata	16560
tggtcttagaa	aggcatgtag	gtaaaaagga	ctcaaaactc	taataatggt	tgggccaata	16620
gtaaaatttg	tagttctact	ccattaaagca	ttcctcaagc	agtgtaaaaa	tcagagttca	16680
agttacactt	tgtgtgttag	atcctttgaa	agccactcta	ccctgtttta	tatgaagcat	16740
ccgcagctaa	aatgaacacc	tagtgaagag	tatgaatgct	gcaatacata	agcagacgtc	16800
agaattgtcc	caagctgatt	ctaagttact	ttaaactatgt	atgcagagtc	agaatatgac	16860
ttacttctta	gaagtaacag	ataattacct	ttggcataat	gaaaaaaact	ttaaattgtaa	16920
gttaatacag	gtattttccc	tttagcaaag	ctttgctttt	aaaagaaaac	ttcaaaactt	16980

aaattaaaat	aggaaatgct	ctactatgta	gtaaaaatac	tttttagatt	actgaagcaa	17040
agaaaaggaa	ggattctatg	agggaggaaa	agtgggagaa	aatgttaaag	aaaaaaagga	17100
agaaggaaaag	aaaagagaaa	aggaggaaaag	aacacaagga	cagaaaggcc	tattgaaata	17160
tattattttct	ttcaaattttt	aaacgagcag	aataaattct	tttgttttat	aactatgaaa	17220
taatctatgt	tcctcttatc	tatgcttgga	aaatttagac	aaaatgttaa	gagtaagtac	17280
tacattggat	ttccgggtct	tcagctctga	aaacaagctg	tttcttaaca	tacgtcaatt	17340
ttctatatatt	catgtcattt	ctatttgcaa	atgttataaa	gttcaatatg	atgtaaaaaca	17400
tggttaaatg	aagttcaaaa	ataagtataa	catacattag	tttggtctatt	ccaaatttca	17460
tgcacattaa	ctcagccaca	catctaacac	agtcagccct	ccctatccag	gggttctgca	17520
tctgcagatt	caactaacca	tgggtcgaaa	atgtttttgt	accaaacatg	tacaggcttt	17580
ttttcttggt	atcattccct	aactacagta	taacaactat	tttcacagtg	tgtacatgtg	17640
tatgaaatat	tataagtaat	ctacagataa	tttaaagtat	acaagagggg	atgcataagg	17700
tatatgcaaa	tactacacca	ttttatatca	gactctcaaa	catcagtaga	atttggtaac	17760
ccaggagggt	cctggaacta	atcaccacga	ggtatcgaca	gatggctata	tataaatcac	17820
tcagtgaatt	caggattcac	attatttcac	aactagtata	attttatgtt	gttcacataa	17880
ttgtgtcaca	acataatacat	gcagacagggt	gactttcatg	aaaagattac	acccaagata	17940
gacatatgggt	ctactcaaat	acggtttcca	aatgtgtatc	caatcttggt	taattataat	18000
caaactcacc	attccattga	taagcgacct	ctaccaacct	gcttatcccc	tccaagcaat	18060
ataacagtgg	ttctctgaac	caatattgac	cctcctttta	attgatagcc	tttttttaaa	18120
aagctaacca	ttgagaagta	catactgttg	aagacagaac	atattctgta	aatgctccc	18180
aagatatcaa	agtcagatga	tacaactgaa	tgtttatgct	agattatatt	tctaagctga	18240
gaattacatt	ttaatatacc	ataagcaatc	tgcaaaagaa	gcaacttgcc	taaagatttc	18300
aggagtttca	agtatgcata	tgtcaatatc	tgtatcaata	tgtaatatca	atataatcaa	18360
tgcacacaac	aatacgtaac	tgtacttata	tcattctcct	agcactaatt	attacaaaca	18420
atctgcatgc	actgcaaagc	aaaagtataa	tataaaatcc	caaaaaacct	tgaaaattta	18480
ataaaaccaa	aaaacaggca	tcacacacaa	gaactgaggc	gtatacttca	ttaatgagta	18540
tgatatcctg	atatgaaatg	tcaaacaaaa	ttaccaggcc	tcagggttaga	aataaagata	18600
ggacattagt	ctttgtattt	ttaaattgat	tttttcttct	aatattcctt	aatgataacc	18660
ctataatatta	cctacttaaa	attatttagca	aatagttatt	ttaaaagtat	gagtaattag	18720
accaaaagca	actctcatat	ttaccacaaa	gaaggaaacca	ctaccaagaa	tcaaagccta	18780
gtaattctgt	tcttaacaga	cagggtgttg	gtattctggc	atgttacatg	aaaatcactt	18840
atgagaagaa	cagaaaaaaa	aattagaagg	tagttttcac	tatggaaata	ggtaagtgat	18900
taagcagatt	ttcttacacc	atgaaattgt	cagcagactc	aataatcacc	ctaaggggca	18960
tcattctgga	tgccgacatt	ctctatgatg	gaaagggact	gaaagtaaaa	tgactaatg	19020
acataaagaa	accaatatcc	aatagtaaaag	ttgaagaaat	aaacattcct	tggaaggtaa	19080
ctaagctgaa	gtttgcaact	accaagaatg	tattatgcca	gcagtaaatt	aggaaactaa	19140
agcccatgtc	aaccaatgaa	aaatgggagg	actgaaatca	atcattaaag	cagcagcaag	19200
gttctaacta	ttctaaggta	taggctacct	ctggcgata	ttatcagagt	tgacaattct	19260
tccaagaaat	tctaacatca	actgtaatct	gaggctcctt	aaaaataat	ataaaccagg	19320
cagtagactt	acattttgta	atattttctt	ctaagagctg	tacatttaaga	ttttatttgt	19380
gatataaata	cttcaaaata	attagctata	gaacgactct	attttcaaca	gttataacat	19440
tttaagccat	ctcacattta	acctaaactt	ttatcaaatg	tcaaaactga	ggccgggtac	19500
ggtggctaac	acctgtagtc	ccagcacttt	gggaggccaa	gatgggcgga	tcacttgagc	19560
ccaggaattc	gagaccaacc	tgggcaacat	ggtgaaaccc	catctctata	aaaaatacaa	19620
aaattagctg	cgccctgggtg	tgtgcgcctg	tagtcccagc	tactagagag	gctgaggggag	19680
gagaatcacc	agggcctggg	agatcaaagc	tgcagtgagc	tgagatcgtg	ccactgcact	19740
ccaccctggg	tgacagagtg	agaccctgtc	tcaaaaaaaa	aaaaaaaaag	aaagaaagaa	19800
aaaaaaatca	aaactgatca	cttgagggtcc	aacttatgtt	tactatatct	acttatattc	19860
ccaaagacat	cttaaggaga	gatgaaatca	taaaaagggtg	aggatgagaa	agaaaatagt	19920
aagtcagtaa	ggtcaatttt	tacatatatt	aggctagcat	aataaaaaata	tgagtgtctt	19980
attattattt	ttttttgaga	cagagtcttg	ctctgttgcc	caggctggag	tgcagtgggtg	20040
caatcatggc	ttactgcaat	gtctgccttc	caggttcaag	caatccttgt	gcctcagcct	20100
cctgagtagc	tgggattaca	ggtgtgcgtc	accctgccc	gtattttcag	gtattttcag	20160
tagagacagg	gtttccacc	gttaaacat	gagttttggc	aggatgggtc	caaactccca	20220
aagtgtctagg	attacatgctg	tgagccactg	cgtctggcct	aaagtgtctt	attataacca	20280
agaattttatt	tgtggagaga	ggtaaagaaa	actcattttt	agtgaataaa	ttaaaactgc	20340
atcattcaca	atctatcttt	caaaatgagg	tattaactat	tttggtcttct	aaaattaccc	20400
catatactac	atgcatgagc	atgggaattg	aagttatttt	attcctaagt	ttgagacttc	20460

atgttttaaat	gtgatcacta	aaaatttccct	aattgatgat	taggaaaata	acttttctgta	20520
aaattccaga	attttagctg	tttcaatctc	ttcatattaa	ggggagaaca	ttatgttttt	20580
acttttctgtg	catgcacttt	ctttattaga	agaaaatgga	ctgagggcag	taagcaaccg	20640
aaaaggaaga	gtaataagaa	gcctgatgtg	tgtgaaaact	ggagaacagt	ctcaaatcat	20700
aaaaagttat	gacagaagag	gcataaaaaa	taaaagtaat	gaacttaata	tatgaaaggt	20760
aataatgatt	aagagcatag	gctataaaagc	cagactggac	tccttggatt	caaactcctgg	20820
ctcttctaata	tactaggtag	gtaaccctga	gcaagtttca	atgaccaatc	tttttctcaa	20880
ttacctcagg	tatataaagg	ggacagtaac	agcatttaac	ccagaggaca	ataaggatta	20940
aataaataca	tgtaaaataa	tttaaaacag	tacctgggtat	tcaataaagc	gcaataaatg	21000
ttagctgcta	ttattattca	tctaaacttt	actttcatta	ccagcaatat	tttttaatat	21060
taaaaatatt	gaataaaaca	atgacctagc	ttagtaaata	aattcataat	gagaaaatgt	21120
tgatttcatt	taataataac	tttagtagtt	tgggataaca	ctttgcatat	tttaatttcc	21180
ccagctataa	ataactcaaa	taatttgcca	tcagatgac	tgttattttg	aagttaacaa	21240
ataaagcatt	tcctaaaaaa	gttctaatac	ataacttttg	ctctcatctt	atgttttaaa	21300
aacaaaatgg	caaactcatc	gcatcaaata	gttccctact	ttataacatg	acaattgttt	21360
taaaatatat	ctgctggaaa	aagcaactga	agtcctagaa	aatagaaatg	taatttttaa	21420
ctattccaat	aaagctggag	gaggaagggg	aaaaacatat	ctgccaaaata	agcttataat	21480
taatagtgtg	tttcagtttt	caaaaatcca	cataggaagc	aatttaagcc	aattttgcct	21540
aagtctcaat	ctcagcgtag	tagatagctt	agggcaatca	aaacttgctg	tggtgggctg	21600
ccccctacag	gactcaattt	acctatttct	tttaaaaggt	gtgtaagtag	gaaatatgat	21660
tcaagtttta	cattaacaat	attaatgcta	aagcagatga	ttatcattca	cgcattoact	21720
ataggaggaa	acagtctctg	agaacctatc	atagagatac	agagagaaat	gaaacaatcc	21780
ttgtccttga	ggaattaata	gtttactgct	tacagagaaa	ctacatacat	ggtgaaatat	21840
ttaaaaatag	ctcatgatat	cctctatgat	attatgtttg	ctatagaaaa	agaacaaggc	21900
tgaagatcta	agatccaagt	tctactgttg	gctctgccat	caaacaataa	gctaaacaat	21960
gtacaagtca	gttttgggga	agctgtctta	ttcccaaaat	gaggagggtta	aattagttaa	22020
ttcttccagc	ctctatggct	ctaattattcc	acagttacat	ttgtcaaaac	aaaaggtaga	22080
aggaaatggt	tcaaaaacag	acttcgcaga	aagaacatct	atatgatatg	aagggtctgg	22140
gcatatgtga	agaaatcaag	gaagacttct	tgaggaaggt	gacatctgaa	gtaacttttag	22200
aagcactctg	ggagccaagg	ctattcccag	gagttaacag	agtcagataa	taaaagatca	22260
aagatgttta	ggggaatagc	atgcagtgtt	atlttggttg	agtcotagcta	tatttttagga	22320
aacatcaaat	taatatcagt	ataaaaactca	acagaatgga	gggagaaaaa	gcaggtagaa	22380
aaatctaaga	accactaaaa	tagttcatct	agaagataaa	ggacccatga	gctaaatcag	22440
tgcaaatggc	aagaagggaa	taaatgaaga	cagttctggg	ccattagaac	tgcaactcaa	22500
caaaagtgat	caaaaagagt	attccaaagt	attgacctgg	taacttgaag	aaaagtaag	22560
aaagaggaaa	ctggacactg	aaacagaaga	agtagattat	gtatttggtg	gtgaatggaa	22620
tagattgggt	gggaccagtt	agaacctcac	agagaagaac	tatgttaaga	ccagaaatac	22680
ggccaggtgc	ggtggctcat	gcctgtaatc	ccagcacttt	gggaggcctg	ggtgggcgga	22740
tcacctgagg	tcaggagtgc	aagaccagcc	tgacaaagat	ggagaaaccc	tgtctccctc	22800
gtctgtacta	atacaaaatt	agccaggtgt	ggtgggtgcat	gcctgtaatc	ccagctactc	22860
aggaggctga	ggtaggagaa	tcgcttgaac	ccgggaggcg	gagggttgag	tgagctgaga	22920
tcgcaccatt	gcactccagg	ctgggcaaaa	agagcgaaac	tcttgtctca	aaaaacaaac	22980
aaacaaaaca	aaacaaaaca	cagaaatata	tcaattaaaa	aagttaggcta	ttcaccagat	23040
atgttccact	ggtcataaaa	caaaagaata	caggaggcat	gacaagccat	catcattgct	23100
gttaaaaataa	ctcacagcaa	aattataatg	atttaagtca	ataacatcta	ataattccag	23160
ctatagtgtg	caatttaatt	tattatgtgc	caggcacaaat	agtttattaa	aggtattacc	23220
tctaattttt	acaataaccc	tattttacag	attataaaat	ggaggcccag	agatgtaagg	23280
tgaacgagcc	aaatcaccta	gttacctgga	atataaactc	agaactgcct	aaatcaaaag	23340
ctctcaatct	taaccacatg	ctatactgat	gcatgtcaaa	gattcaattc	attcagattt	23400
ttcaaggtta	tcggaaaacc	tatgtagata	aaaatttcca	aaataatcaa	ggatatgtaa	23460
cttttacaga	aagcaatcac	tgatcatcta	ttgcaatact	catgttctta	agcaatatac	23520
tgagttgaaa	tttttatatt	ttataaataa	ttagaaagaa	tacatttttt	aaaactttta	23580
aaaacacctc	agttttttat	ctcttcccca	aatttcaaca	aaatccattt	atccaaactt	23640
gaggttgaa	cattaaagtg	gtgatatcat	cagtaataagc	agagttagga	cctgaatat	23700
actctcctcc	ataaaagcaa	caagaacaca	aaaattctca	aaatgaactt	tttctgaaat	23760
ctttcaaaag	ccccactctc	agaaaactgt	cattatttga	tctgccagtt	ccctagaaaa	23820
acctccctca	taggacatta	tttgacttga	ctcagagctc	actcagtgca	aacaatttta	23880
tcaccaggag	agttttgtgga	aaatcagtg	caattgttaa	acatcacatc	tgccatgaga	23940

tagcaataac	agatgggaca	aacaagctaa	ccaaaaaatt	aaaagaaaaa	cctggggaat	24000
aagaaatcca	aaggggggtct	gaaaagttct	aacatatcttc	tgataatcca	gaaagccata	24060
cacatgtata	gagctgtgta	cacgctcaaa	aaacatctac	gaaggcccta	aactctcacc	24120
tatgggaaac	cctgaggctc	tgtacaagaa	gaaagtaaaa	tccagttata	aattgcttgc	24180
cgtatcattg	aaggcaatgc	cccaacattc	acacataggc	ccctggcaaa	gattggaaga	24240
tactctagtt	ctaggcattc	aagaaaatct	cttctaatac	tcagatgata	actaaactca	24300
ccaagcagta	acttttagggg	cctgtgtgat	aaaaaataaa	aacctgaaag	aattagttca	24360
ggaaagaaac	taaacaagca	acagcaacaa	caaaaacaga	ccttgggaaa	ggggggaagc	24420
atctggtttc	cagagttatt	ctgttatact	atataaaata	ttcaggtctc	aacaacaaca	24480
aaattacaaa	gacatgcaaa	gaaacaagta	taagccacaa	actgggggga	aaaagcagca	24540
gaaactggcc	ctgaaaaaga	ccagatgctg	gacttactgg	acaaagactt	taagagagt	24600
atttttaata	tgcgcaaa	actaaaaaaa	agtttatcta	aagaactaca	ggaaagtata	24660
agaacaatat	ttctgatcct	tcagaagaac	cactttttgt	cactacagat	tagttctgtc	24720
tgggtctagaa	cttcttaaaa	acagaatcat	agagtataat	ctctttatat	cagctctttt	24780
tactcaacac	aatgttgtgt	gagattttatc	catgttgttg	catgtatcat	tcccaaacag	24840
aaatagaaat	tatagagata	aataggagtt	acaaaaaagt	accaaacaaa	aattctggag	24900
ttgaaaagca	caaaaactga	attaacttga	ggggctcaac	agctgatttg	ggcagccaga	24960
agaatgaatc	agcaaatcta	aagataggctc	aattgcgaga	aagagaggga	agaaggagg	25020
aagggaaggaa	aggaggctca	gagacccaag	agacaccatc	aggcatacca	atatacatat	25080
aatgagaggc	ccagaagaag	atgcagaaaa	agggtcagag	tatctgaaaa	aataatggcc	25140
ctaaacttcc	cgaacttgac	cccaaaaatt	aatctacaca	tccaagaaga	taaacaaact	25200
aaaaagaata	aatcaaaagc	gatccacacc	taggtacatc	ataatcaa	gactgaaata	25260
taaagagaga	ctctcaaaac	aggcaaggga	cttatgtaca	aaacatcttc	agattaataa	25320
caaatttctc	atcagaaatg	atgttgtcaa	taggcaatca	gatgacataa	tcaaagcact	25380
gaaagaagta	gaatgtctgg	gacctggaat	gctgggtggac	acctgtaatc	tcagtatttt	25440
gggtggccaa	ggtgggagga	tcacttgagg	caaggagttg	aagaccagcc	tgggcagcag	25500
aaagaggctc	tgtctctaca	aagaataaaa	agattggctg	aatgtgggtg	tgtggacctg	25560
tagtcccagc	tactcaggcg	gctaagggtg	aaagatcgct	tgagcccagg	agttggaggc	25620
tgcagtgtgc	tatgactgtg	ccactgcact	cttgtagtgg	agaccctgtc	tctataaaga	25680
aaaaatgtca	acaaaaaact	acatgcagaa	aaactgcact	tcaagaaatg	atcagtacct	25740
tgaagctctg	aaggtgctta	agactgtaga	tcaataccat	agaaaaataat	ttagtattta	25800
ggaatgtaag	aaaatttaaga	cagccttggt	tgataactac	acataatact	gtaactgttc	25860
ttgcactggt	ctgggtattg	tcaagctatg	agcacaaaact	gatgactgaa	atacagaata	25920
cagaacagga	tataaaatct	tatcaggtaa	agttaggcaa	gcaattacta	gttgtaattc	25980
aaacttgaag	agaaggaaata	aggaaccaac	tcaaacagg	cagcaatgaa	ttgtaaaaaa	26040
gcttaaggta	aaacaaacag	ggaaataaaa	caactcagaa	cctaagcata	tcgtaagaac	26100
ctaacttaac	aaggaggggc	ttaaactgat	tattttacag	cttgggtgca	attatcccac	26160
aaaaaacttt	caggagtttc	accagtccat	aaactatttg	gttattagaa	aatagcttta	26220
ttgggctacc	ctctttgggt	ccctccctt	tgtatgggag	ctctgttttc	actctattaa	26280
atcttgcaac	tgcactcttc	tggtccgtgt	ttgttacggc	tcgagctgag	cttcactct	26340
ccatccacca	ctgctgtttg	ccgccatcgc	aggcctgcc	ctgacttcca	tccctctgga	26400
tctagcaggg	tgtccgttgt	gctcctgatc	cagtgaagac	cccattgccg	atcccagctg	26460
ggctaaagac	ttgccattgt	tcctacgcgg	ctaagtgcgc	gggttcaccc	taattgagct	26520
gaacactagt	cactgggttc	cacggttctc	ttctgtgacc	cgtggcttct	aatagagcta	26580
taacactcac	cgcgtggccc	aagattccat	ttattggaat	ccatgaggcc	aagaacccca	26640
ggtcagagaa	cacgaggctt	gccatcatct	tagaaccagc	ccgccaccat	cttcggagtt	26700
ctgggagcaa	ggacccctcg	gtaacaattt	ggcgaccaca	aagggaacctg	aaccgcgaac	26760
catgaaggga	tctccaaagc	ggtaaatattg	gaccactttt	gcttgctact	ctggcctatc	26820
ccttagaatt	ggaggaaaat	actgggcacc	tgtcggccgg	ttaaaaacga	ttagcatggc	26880
cgccagactt	tagactcagg	tatgaggcta	tctggggaag	ggctttctaa	caaccctcaa	26940
cccttctggg	ttgggaacct	tggtctgcct	ggagccagct	tccactttca	attttctctg	27000
ggaagccaag	ggctgactag	aggcagaaa	cttcgttccc	gaactcccgg	cattagccgg	27060
ttgagatcat	gtcgcagcca	gaagtctcta	ctcaacagtc	gcccattcgt	gcgctcctac	27120
cttcccttct	gtccacaccc	tcctgggtcc	caaccacgac	tttcttgaaa	gtgtagcccc	27180
aaaattctcc	ttacctctga	atctacttcc	tctgatccct	gcctcctagg	tactaatggg	27240
tgagactttc	atttccctcta	gcaagttgta	tctccaaagg	gatctaagga	agctctatgc	27300
tgcgccctta	ggcatctagg	ctataaaccc	agggagtctt	gtccctgggtg	tccctcctga	27360
tttaggtata	cagctctaga	catgggcagt	tatgtgggac	ctgttcccca	ccacccttgc	27420

cagggcccca	agtttgtaaa	tggctaagag	aggaaacaga	gagagacaga	gagaaagaga	27480
cagtggagaga	cagacagaga	cagagagaga	gagagacaga	gaggagagag	agagagacag	27540
ggaggacagg	gagagagaca	gagaggagag	ggagagagac	aaagaggaga	aagaggcaga	27600
gagacaaaca	gggagtcaga	gaaagaaaaga	caaagataga	aatagtaaaa	aaaaacagtg	27660
tgccctattc	ctttaaaagc	cagggtaaat	gtaaaaccta	taattgataa	ttgaaggctc	27720
tctccgagac	cctataaacac	tccaatacta	ccttgttgtc	agcgtaaaaca	agggcgtagc	27780
ctgaaaacac	taagaccact	gacaacccat	agccttctca	tcaaaaatcc	ttaacatcca	27840
gtgacctgag	gatggcccaa	atgcattcaa	tctgtagcgg	caactgcttt	gctaacagaa	27900
aaaagtagaa	aagtaacttt	tagaggaaac	ctcattgtga	gcacacctca	ccggttcaga	27960
attattctaa	gtcaaaaaag	caaaaaggta	gcttattaac	tcaaaaatat	taaagtattg	28020
ggctattctg	tcagaaaaag	gtaatttaac	actaaccact	gataattccc	ttaacccctg	28080
agatttccct	acaggggatt	taaatcttaa	ttaccataca	aagggtccgac	cagacctagg	28140
aggaactccc	ttcaggacag	gatgatagat	ggttcctccc	aaatgactga	ggaaaaaacc	28200
acaatgggta	ttcagtaatt	gatagggaga	ctctgttgga	agcagagtta	gaaaaattgc	28260
ctaataattg	gtctcctcaa	atgtcagagc	tgtttgcaact	cagccaagcc	ttaacgtact	28320
taccgaatca	aaaagactat	ctcaatcctg	actcaaaagc	ttacttatac	cctctctgaa	28380
acgaatttgc	ctaagaactg	ttgtttatgg	gaatgcactc	tgatggagca	gctgggtgtg	28440
tatgaaatac	tcaggaactc	agcctagctc	taggactcac	ccctgagcac	aaaggcaatg	28500
ttgggcacgc	tggtaaagga	ccactagaat	ccagcagccc	ggaccccttt	ctttgtgatc	28560
aagaaaggcg	ggaaaagggg	tgagggctgc	tacatcagtg	agcataacta	atccgataag	28620
cagaggtcca	tgggtggtta	cacaccccg	aaaggaataa	gcattaggac	catagaggac	28680
gctctaggac	taatgctcat	cggaaaatga	ctagtgggtc	tggtatccct	atgttctttt	28740
ttcagatagg	aaacggtccc	ctcaaggcaa	aaacaccctc	aagatgtatt	ctggagaatt	28800
gggaccaatt	tgactctcag	atgctaagaa	aaaaaagaca	tattcttctg	cagtaccgcc	28860
tggcaacgat	atactcttta	agggggagaa	acctggcatc	ctgagggaa	cataaattat	28920
aacaccatct	tacagctaga	cctcttttgt	agaaaagaag	gcaaatgggt	tgaagtgtca	28980
tacgtacaaa	ctttcttttc	attaagagac	aactcgcaat	tatgtaaaaa	gtgtgattta	29040
tgccctacag	gaagccctca	gagtctacct	ccctacccca	gcacccccc	gactccttcc	29100
ccaaataata	aggacccccc	ttcaacccaa	acggtccaaa	aggagataga	caaagggtgt	29160
aacaactaac	caaagaatgc	caatattccc	cgattatgcc	ccctccaagc	ggtgggagga	29220
gaattcggcc	cagccagagt	gcacgtacct	ttttctctct	cagactttta	attaaaaatg	29280
acctaggtaa	attctcagat	aaccctaatt	gctatattga	tgttttacaa	ggtttaggac	29340
aatcctttga	tctgatattg	agagatataa	tgttactgct	aaatcagaca	ctaaccctca	29400
atgacagaag	tgctgcctga	actgcagcct	gagagtttgg	cgatctctgg	tatctcagtc	29460
aggtcaatga	taggtcgaca	acagaggaaa	gagaacgatt	ccccacaggc	cagcaggcag	29520
ttcccagtg	agaccctcac	tgggacacag	aatcagaaca	tggagattgg	tgccgcagac	29580
atttgctaac	ttgctgtcta	gaaggactaa	ggaaaactag	aaagaagcct	gtgagttatt	29640
caatgatgtc	cactataaca	cagggaaagg	aagaaaatcc	taccgccttt	ctggagtgtc	29700
taacggaggc	attgaggaag	catacctctc	tctgtcaact	gactctactg	aaggccaact	29760
aatcttaaa	gataagttta	tcactcagtc	agctacagac	attaggaaaa	aacttcaaaa	29820
gtctgcctta	ggcccgaac	aaaacttaga	aaccctattg	aacttggaac	cctcagtttt	29880
ttataataga	gatcaggatg	agcaggcaga	atgggacaaa	tgggataaaa	aaaaggccac	29940
cgcttttagtc	atggccctca	ggcaagcgga	ctttggaggc	actggaaaag	ggaaaagcta	30000
ggcaaatcaa	atgcctaata	gggtttgctt	ccagtgcggt	ctacaaggac	actttaaaaa	30060
agattgtcca	aatagaaata	agccgcccc	tcgtccatgc	acctcgtgtc	aagggaatca	30120
ctgtaaggcc	cactgcccc	ggggacgtag	gtcctctgag	tcagaagcca	ctaaccagat	30180
gatccagcag	caggactgag	agtgcctggg	gcaagcacca	gcccattgcca	tcaccctcac	30240
agagccctgg	gtatgcttga	ccattgacgg	ccaggaggct	aactgtctcc	tggacactgg	30300
tgtggccttc	tcagtcttat	tttctgttcc	cagacaacgg	tcctccagag	ctgtcactat	30360
ccaaggggtc	ctaggacagc	cagtcactag	atacttctcc	cagccactaa	gttgtgactg	30420
gggaacttca	ctctttttcac	atgcttttct	aattatgcct	gaaagcccaa	ctcccttgtt	30480
agggagagac	attctagcaa	aagcaggggc	cattatacac	ctgaacatag	gagaacaccc	30540
gtttgttgtc	ccctgcttga	ggaaggaatt	aatcttgaag	actgggcaac	agaagaccaa	30600
tatggacgag	caaagaatgc	ccgtcctgtt	caagttaaac	taaaggattc	tgccctcttt	30660
ccccaccaa	ggcagtaccc	ccttagaccc	gaggctcaac	aaggactcca	aaagattaag	30720
gacctaaaag	cccaaggcct	agtaaaagca	tgcaatagcc	cctacaataa	tccaacttta	30780
ggagtacaga	aaccagtg	acagtgagg	ttagtgaag	atctcaggat	tatcaatgag	30840
gtcactgtcc	ctctatacct	agctgtacct	aacccttata	ttctgcttcc	ccaaatacca	30900

gaggaagcag	agtgggtttac	agacctggac	cttaagggatg	cotTTTTctg	catccctgta	30960
catcctgact	ctcaattctt	atttgccttt	gaagatcctt	caaaccctaat	gtctcaactc	31020
acctggactg	tttcacccca	agggttcagg	gatagccccc	atctatTTtg	ccaggcatta	31080
gcccaagact	tgagccgggt	ctcatacctg	ggcactcttg	tccttttggt	tgtggatgat	31140
ttttactttt	agccgcagct	tcagaaacct	tgtgccatca	agtcacccaa	gtgctcttaa	31200
attttctcgc	tacctgtggc	tacaaggttt	ccaaaccctaa	ggctcagctc	tgctcacagc	31260
aggttaaata	cttagggcta	aaattatcca	aaggcaccag	ggccctcagt	gcctattctg	31320
gcttatcctc	atcccaaaac	cctaaagcaa	ctaagaggat	tccttgacat	aacagggtttc	31380
tgccaaatat	ggattcccag	gtacggcgaa	atagccagac	cattatatac	actaattaag	31440
gaaactcaga	aagccaatac	ccattttagt	agatggacac	ctgaagcaga	agcggctttc	31500
caggccctaa	agaaggccct	aacccaagcc	ccagtgttta	gcttgccaac	ggggcaagac	31560
ttttctttac	atgtcacaga	aaaaaacaga	aatagctcta	ggagtcctta	cacaggtcga	31620
tgagcttgca	acccatggca	tacctgagta	aggaaattga	tgtagtggca	aagggttggc	31680
ctcattgttt	atgggtagt	gcggcagtag	cagtcttagt	atctgaagca	gttaaaataa	31740
tacaaggaag	agatctgtgt	agacatctca	taacgtgaac	ggcatactca	ctgctaaagg	31800
agacttggtg	ctgtcagaca	accgtgagga	aagtaactaa	aatcgtaaat	ccccatggcc	31860
ctcccttate	atatttttct	ctttactgtt	ctcttaccct	ctttcactct	cactgcacc	31920
ccctcatgct	gctgtacaac	cagcagctcc	ccttaccctaa	agtttctatg	aagaatgcgg	31980
cttcccagaa	atattgatgc	cccatcaaat	aggagtttac	ctaaaggaaa	ctccaccttc	32040
actgcccaca	cccatatgcc	ccacaactgc	tataactctg	ccactctttg	catgcatgca	32100
aatactcatt	attggacagg	gaaaatgatt	aatcctagtt	gtcctggaag	acttggagcc	32160
actgtctgtc	ggacttaact	caccataact	ggtatgtctg	aggggggtgg	agttcaagat	32220
caggcaagag	aaaaacatgt	aaaggaagta	acctcccaac	tgaccgggtg	acatagcacc	32280
cctagcccct	acaaaggact	agatctctta	aaactacatg	aaaccctcca	taccataact	32340
tgcttggtaa	gcctatttta	taccaccttc	actgggtctc	atgaggtctc	ggcccaaaac	32400
cctactaact	gttggatgtg	cctccccctg	tatttcaggc	catgcatttc	aatccctgta	32460
cctgaacaat	ggaacaacta	cagcacagaa	ataaacacca	cttccgtttt	agtaggacct	32520
cttgtttcca	atctggaaat	aaccataacc	tcaaacctca	cctgtgtaaa	atttagcaat	32580
actgtagaca	caaccaactc	ccaatgcate	agggtgggtaa	ctcctccac	acgaatagtc	32640
tgccctaccct	cagggaatatt	ttttgtctgt	ggtaccttag	cctatcgttg	tttgaatggc	32700
tcttcagaat	ctatgtgctt	cctctcattc	ttagtgcctc	catgaccatt	tacactgaac	32760
aagatttata	caattatgtt	gtacctaaag	cccacaacaa	aagagtactc	attcttccct	32820
ttgttatcgg	agcaggagtg	ctaggtggac	taggttctgg	cattggcggt	accacaacct	32880
ctactcagtt	ctactacaaa	ctatctcaag	aactcaatgg	tgacatggaa	tgggttgccg	32940
actccctggg	caccttgcaa	gatcaactta	acttctagc	atcagtagtc	ctccaaaatt	33000
gaagagcttt	agacttgcta	acotctgaaa	aggggggaag	ctgtttattt	ttaggggaag	33060
aatgttggtt	ttatgttatt	ttagcggaag	aatgttggtt	ttatgttaat	caatcctgaa	33120
ttgtcacaga	gaaagttgaa	gaaattcgag	attgaataca	acgtagaaca	gaggagcttc	33180
aaaaacacca	gaccctgggg	cctcctcagc	caatggatgc	cctggattct	cccttcttta	33240
ggatctctag	cagctctaatt	attgatactc	ctctttggac	cctgtatctt	taacctcctt	33300
gttaagtttg	tctcttccag	aatcaaagtt	gtaaagctac	aaatcgttct	tcaaatggaa	33360
ccccagatga	agtcocatgac	taagatctac	cgtggacccc	tggaccggcc	tactagccca	33420
tgctccaatt	gtaatgatat	cgaacgcacc	cctcccgagg	aaatctcaac	tgcaacaacc	33480
ctactatgcc	ccaattccgc	aggaagcagt	tagactgggc	gtcagccaac	ctccccaaca	33540
gcacttgggt	tttccctgtg	agtgggggga	ctgagagaca	ggattagctg	gatttccctag	33600
gccgactaag	aatcccaaa	cctagctggg	aagggtgacca	catccacctt	taaacactgg	33660
gcttgcaact	tagctcacac	ccgaccaatc	aggtagtaaa	gagagctcac	taaaatgcta	33720
attagacaaa	aacaggaggt	aaaaaaatag	ccaatcatct	atcgctgag	agcacagcgg	33780
gaaggacaat	gatcgggata	taaaccaggg	cattcaagcc	ggcaacggct	accttctttg	33840
ggccccctcc	ctttgtatgg	gagctctctc	tgtcttcaact	ctattaaata	ttgcaactgc	33900
aaaaaaaaaa	tagcttaatt	gaagaataaa	ttaatacaat	aaaaggaata	catttttaagt	33960
atacagttca	aactgtaaca	gtgttacagt	ttcaagagga	ccccctcaac	aagatattgg	34020
gcatttccat	catgcccata	aagttccctc	tgtccctta	ctggttgggt	ccatctctac	34080
tacaccctcc	tgacctggcc	cagaccttgg	cctcagaaga	atcatttttt	tgtcactaca	34140
tattagtttt	gtctgttcta	gaacttctta	aaaacagaat	catagagtat	gttctctttg	34200
tattggttct	ttttactcaa	tgtaatgttc	tgtgacattt	atccatatta	ttgcatgtat	34260
tattcccttt	aatcctgaat	agtatgctgt	tttaggaata	taatgcaatt	gtttattcat	34320
ttacctgttg	acagatatct	gagctattat	gatggatatt	atgaataatt	ctgctatgaa	34380

cactttctgta	caatgttttc	toggacatat	attttccattt	ttcttgagtg	gagctgttag	34440
aactgtttgga	tcagaaagta	agcatatgtt	gaattttgaa	agaaactggg	aaactcttgc	34500
ctaaagtgat	ttgtaccatt	ttacactcct	actaataatg	tatgagagtt	atatttgcct	34560
cacagccttt	ttactacttt	gttaatcttt	ttagtactgt	caaccttttt	aattttatcca	34620
atctagggaa	cgtgaagtag	tatctcactg	ttattttcat	tttctgatg	agtaacaata	34680
tcgtgtatct	tttcatgtgc	ttattagcca	ttcctatatc	ttttgtgaaa	tagtttaact	34740
aaatttgtaa	ctaaagggtg	tttctgagt	ttcaggtagt	aagcctattt	ccctcaagtg	34800
aataaactac	agtcttgga	tgaaaaatta	aacacagtg	agacattttt	tgtataagtt	34860
gttttactct	gtgtatgtct	ggtttgctta	gtctattatt	atatgcccc	tgaaagcaaa	34920
cacagtgcct	atttcactaa	tgagtatcac	tagcacatag	aactgtgcct	gccccaaagca	34980
tgaactcaat	aaatatgtta	atgtgtatgc	atgcacatac	atctacatgc	atgtacatct	35040
atacacacat	ataaacatat	atlaattttt	agaccacaaa	atctaagaaa	actaattctt	35100
gagcctctgg	tttgaagaat	tctcaaatta	ttacatatc	tttatgttcc	actccacatc	35160
cactgtacct	gaaatagccc	tactgttcta	ctttggtaaa	tcaggcaaat	tttaattttt	35220
aaataattaa	gattccaact	aatttttaaa	tataatttga	aagttaacaa	tgaaatacat	35280
tacataaaaa	gaaaatttta	aataaaagca	aaactaaacc	caataagagg	aaagaaagtt	35340
gggctgtatt	tctttaatcc	tttaaaattc	aaatcacaca	atgctccaat	gaaatcttca	35400
tttaactgaac	caaactatgc	ccatgaaaga	tctcatatgc	aactgctaaa	acctcaataa	35460
acatatctcat	cttcttgcaa	aaaagatat	tctttataat	atgcacatgc	agtatatact	35520
attttgaggc	agatttgtac	tttagtctct	gttccattgc	ttaccggctg	gctgtccttt	35580
gtctggtcat	tgacctccaa	cttaaaaaat	aatacttgcc	ttgtctaccc	cacagaagtg	35640
ttatgaaagt	caaacaagg	agcataaag	tattttaca	gatataaagt	gctataatac	35700
agatttttaa	aatcactcta	catcccataa	tactttgttg	tacaatttta	gagcaatagt	35760
agaaaataac	aattattgcc	taattgaaaa	tccagtcctg	aattccataa	aatgtatgat	35820
atgaacatta	tagtacatca	tattacgagc	cccaaataat	cactgcttat	atagttgggt	35880
aggatttct	tagtttggtc	atatagttta	tatatttatg	cagtccctat	tttgtgagag	35940
gcattgtgag	gagcataaag	acataagcac	agtacagagc	cttagcttct	ctacattttac	36000
taaagaagac	ttcttcttgg	gtattttaat	aatattttaa	gtattctggg	aagaaatgaa	36060
attaacttca	tagactgacc	ttagattact	atcattacaa	aaagatgcct	gagtgatctg	36120
tctttaacat	accagtat	atcttataac	tgttatat	acttgaatca	gaagtgaagt	36180
ccttttaagc	actaagcatc	cattctatac	tttcttgtct	ttacatatga	gatacaaatc	36240
atatttttaa	aactttttatt	tactttttatt	ttttagagac	ggagtcttgc	tctgtagccc	36300
aggctggagt	acagtggcat	gatcttggct	caccacaatc	tccacctcca	cttcccagg	36360
ttcaagtga	caaatcatac	ttttaagcac	agattctcaa	catgtatcct	agcatgctac	36420
tgccataact	agggtgtgaa	ttaagtatta	aagacagctt	accccaaata	ttactgtaac	36480
atatatctct	aaatgaaaaa	gaacatat	acaactatac	ttggatggga	ttctgggagc	36540
taacccatcc	ctctctcccc	tttctcccaa	attccatctc	ctattaacac	accagctctc	36600
ctgagctaag	cagctcctgg	gggtggggaa	gggtgtacat	ggagaaagct	agaacctcta	36660
cagtgttttc	ctctctggga	ggaactagca	ggcatacgaa	cagaaaaagc	tgataaaaag	36720
gctgaatcct	ttctattcct	gaggcagaca	gagagaagac	cagggaaaca	agagacttgc	36780
accaagagcc	ctgccaggta	ttgatacctt	tgatactgag	aaaatatctg	ggatatgaaa	36840
tacaaatgct	aaataagtat	ctttgaaata	gggtgaaaag	aataaagggt	cttgatgagt	36900
aaaatgggta	gtatttttta	ataacctgat	aatgagcttt	aggaaaagg	aaggtcaacg	36960
ttatggaatg	aaaacacaga	ggtaccaa	ttaaaagcat	aaaaaaaagt	ggaggggggg	37020
aacccaataa	cttcatcaaa	ctagcaata	acttagtatc	atttctaat	agaaacgcta	37080
gaaggaaatc	acttagatct	gataaagact	aggctataat	tctaactgat	gaaacactta	37140
aactgtatca	attaatacca	gaaaacaaac	acagaaaagt	ctactagaac	catcattatt	37200
cagcacagtc	ttggtaatgc	aatactataa	tagcaatgca	ataaagcaag	aaaaaaaaaa	37260
gtttgtaaaa	acacaatagg	atgagatttt	tgtttttcca	atgccataaa	taactagaaa	37320
tggaacaaa	ataaagaaaa	acaaaatcta	caaaacacct	ggaaataaaa	agaaaaatgg	37380
tctatttgaa	gaaaacctta	aaatctatgc	agaacataaa	acaaaatctg	aataaaaaaga	37440
aatatcatgt	tcttgtctgg	gaagacttaa	tatcataaga	aagtgaatta	tatcaaaatt	37500
ttaaatcgaaa	tttaatgtat	ttccatctct	aatcacagac	gacactatgg	ggaactgaat	37560
aagtgatttt	aaaagtcag	gaaaattaat	aactgagaat	aacctgaaa	agtatgaaaa	37620
aaggagacaa	atgaattgct	ccaacagata	tcagaacgct	aaaattaaat	aaaaatacta	37680
ctaggataag	aaaatacata	tactgatgta	atgaataaag	aatccagaat	tagattccag	37740
taagtcaaac	tactttacta	taaaccagg	gtggcatatt	catccagtg	gaaaaggaca	37800
gtaagaagt	agtaaaactat	ggccactgg	ccaaattgtg	gcctctgcct	atttttgcaa	37860

ataaagtttt	actgggacaa	agccaagcct	atcatttgca	aattgtctat	aaatatatttc	37920
atgttacaga	atcacacagt	ttcaacagag	accatcttgt	ctacaaagct	gaaaatatct	37980
actatctggc	ccttgaagaa	agtttgccaa	accttagttt	atataataaa	agatcagcta	38040
tctcatagac	acctatctca	cacaacacat	tgtgggaaag	gaccttcttt	tttttttgag	38100
acggggtctt	gctctgttga	ccaggetgga	ctgtagtggc	atgatcatgg	ctcactgcag	38160
cctcaacctc	ccaggttcaa	gtaatgctcc	caccacagaa	tcccaaacag	ctgggagaga	38220
tgtgtgccac	tacgcctggc	taaggggcct	ttttaacaga	gaaagaaatc	cacatactac	38280
taagaaaaag	aagggcatat	ttgatataata	tttataattt	ttatatagat	atcataaaaa	38340
tcaagatgaa	ttatacagtt	atatttttgca	atgtgtttga	cggtaaaagt	ttaatatcta	38400
taaaaattat	tttataaaaat	atctttaata	tatttataga	tattataata	taaaatatct	38460
ataaaaattat	tttataaaaat	aaaaagttaa	gaagaaaaga	taggcaaaac	aaaatacagt	38520
gcaattttaca	gaaaaccaag	tccaaatggt	caacaaagat	aaaacagatt	tataaactca	38580
ctaagtgtga	gagaattatt	agttaaagta	aaaatatctc	tctataccca	caatactact	38640
aaaaatcaga	gttataatgc	cctattgctg	gtggagatgt	aaggggagaa	gcatgctctc	38700
atatactggt	agtgaaaatt	taaactaata	cattttttgaa	aagtaagctg	gcaatttttt	38760
ttttaatctc	tacctttttga	tgcaaaaact	cattttttggg	tacctattcc	ataccttaaa	38820
aaaaatacat	atgcttactg	tagtactggt	tataatggta	aaaactagaa	aaaaagaaaa	38880
cttgatagtg	aatactgaac	aaattacagt	gcattctacag	attaaacata	atgcagccat	38940
taaaaaagaa	taaattaggt	tgggtgcggt	ggctcatgcc	cgtaatccca	gcactttggg	39000
aggccaaagc	aggcggatca	cttgaggcca	ggagtctcag	accagcctgg	ccaacatggc	39060
aaaaccctgg	ctctacaaaa	aatacaaaaa	ttagtgcggc	atggtggtgg	gcacctgtag	39120
tcccagctac	tcaggaggct	gaggcaggag	aatcacttga	gcctgggaga	cagagattgc	39180
agtgaagcaa	gatcatgcca	cagcattcca	gtccagggtga	cagaacgaga	ctctgtctca	39240
acaaaaagaa	caaattaaac	cctacaactc	atcaacaaaa	atacccaaac	ccaattcaaa	39300
aatgggcaaa	ggacttgaat	agacatttct	tcaaggatga	taaacaagca	catgaaaaga	39360
tgacagcac	tattcattag	tgattacatc	ccacatgcat	taggatggct	agtatgaaga	39420
acagaaaata	ataaatattg	gtgaagatct	gaaaaacaga	aacctttgtg	cactgtttgt	39480
gggaatgtaa	agtgggtacag	ctactacgga	aaacagtatg	gccatttctc	aagaaaaata	39540
aaataaaaatt	atcttatgat	aggaatatgc	attttctggg	aaatacccca	aataactgaa	39600
aacagggtgt	acacccattt	caacattttac	atgtcaattc	aactgggcca	gaatacccag	39660
atatttgttc	aaatatattc	ctggatgctt	ctatatatat	gtttttttggc	tgaggttaac	39720
atttaaatgt	gtggattctg	agtacagcag	attaccatcc	acaatgtagg	tgggcctcat	39780
ctactcagtt	gaaggctctta	cagaaaaaga	ctgacctccc	ttgagcaaga	aagaattcag	39840
gcaacagact	gccttttgga	tcaactgcaa	ctcttctctg	agtcaacagc	ccatcccac	39900
accctggctt	gggtgagtcca	gggtctgatg	aggtaggctg	cagactcaag	gaagagctgc	39960
caaaaccagg	aaagccaatt	cattaaaaata	aatctctctc	tacacaaaaca	cacacacaca	40020
ctaccaccac	caccatgatg	gttctgtttc	tctggagaat	gctaatacac	ccctgttcat	40080
ggcagcatta	ttcacaatag	ccaaaagggtg	gaagcaactc	cagcagatga	atggagaagc	40140
aaaatgtggt	atgtatatata	aatggaatat	tattaagcct	ttaaaaagtg	gaaattatat	40200
ctatctatat	ctatacacac	atactcacac	acacacacac	acatttatag	aagcaggggt	40260
ttcaccatgt	tgtcaaggct	gggtctcgaac	tcttgggctc	aagcaaaaccg	cctgcctcag	40320
cttcccaaaag	tgctgagatt	acatgtgtga	gccaccacac	ccagccaaaa	aaaggacatt	40380
ctgacacata	atacaatata	gataaacaat	gaggacatca	tgatatgcga	aataagcctg	40440
tcacaaaaag	gcaatttagtg	tatgattcct	cttgatagag	gtacctatgg	atgtcaaatc	40500
cataaagtag	aatggggaaa	cagagagttg	tttaatgggt	atagagtttg	ttttgcaaga	40560
agaaaagagt	tttgagagaat	gaatgtacaa	cagtgtgaac	ataattaaca	ctactgaaaa	40620
tggttaagat	tataaatttt	atgttacatt	tatttttacc	tgattaaaaa	ttaaaacaaa	40680
ataatattaa	ggaaaaatac	tataaataac	aacaacaaaa	aaaacacctc	aagcaactta	40740
cattcacctg	ggaaacagaa	tacatcctat	tctgctagag	atatatctgc	agttcaaaat	40800
ttattacaaa	tgatgtttgtg	tatctttttg	aatgactga	aaaactaaat	taaaagcaat	40860
aatattcagt	ttactaacca	gtaagtcctt	ctttcatggg	tcctgacttt	tctgtaagat	40920
gttattgcaa	gatattctact	aaaatggaaa	acaactgaaa	aggcaaaatt	ataatttctt	40980
atcaacatcg	ctaaaaccct	ggaggggaag	aatcctaaca	aacatggcca	taatttgcca	41040
catatttcta	ctgtcctcac	ttttcaaaat	ccagaaatca	acatttctgg	aaacaaaaa	41100
gagtctaaaa	tttggtcctc	tcttcagttt	agaaggtgcc	aagttaatcc	ctgacatcct	41160
agttttccatt	ttcaaaaatg	tactttttct	ctcccaaac	cggtatctag	attcttaaat	41220
attttttagca	catagaagtt	aaatagattt	gcttaaccaa	aatagccagt	aaacctccca	41280
aaagaattaa	aatattaatg	gcgctttaat	gatacaaatg	aacaacttta	cattcaatcg	41340

tcaatgggaa	aggaagcaga	attctgagga	ttatgaaagt	aaacaaaacg	aagttcaaat	41400
tctactttat	tttacttttt	tgtaactaat	gaacaacttc	ttccaaagac	aagtaggaaa	41460
tacaaaaatt	agccaggcat	ggcacatgcc	tgtagtccctg	gttacttgga	aggctgaagt	41520
gggtggatcg	cttgagcccg	gaaggcagag	gctgtagtga	gctgagatca	catcacgca	41580
ctcaagcctg	ggtgacagag	caagaccctc	tctggggaaa	aaaaaaaaaa	aaataggctg	41640
ggcgagtg	ctcacacttg	taattccagc	actttgggag	gctgaggcag	gtgggtccac	41700
tgaggtcagg	agttctagac	cagcctgacc	aatatgggtga	aacctgtct	ctactaaaaa	41760
tacaaaaatt	agccaggcat	ggtggtgggc	aattgtaatc	ctagctactc	gggaggctga	41820
ggcaggaaaa	tcgcctgaac	ccaagaggcg	gagggtttcag	tgagccgaga	ttgcactagt	41880
gcactccagc	ctgggcgaca	gagcaagact	tcactctcaa	ataaataaat	aagtaagtaa	41940
ataaaattaa	aaaatatata	aaaataaaac	aaagataaagt	aggaaccatc	cttttttttt	42000
tttttttttt	ttttttttta	agatagggtc	tgtttctgat	gcccaggctt	gagtgtagtg	42060
gcatgatcat	ggctcactgc	aaccttgacc	tctcaaatac	aagtgactct	cctacctcag	42120
cctcccaagt	agctgggact	acagggtgctt	accaccccat	ccggctcatt	taaaaaaat	42180
tttttgtaga	ggtggggctc	cactatgttg	tatccaggct	ggtctcattt	taactttat	42240
agaaaacaag	cattgtttta	tcagcttctt	gtttttttta	aactaaaaat	aacactgcta	42300
ggttgtttct	atgaagattc	tctaaattta	tttataacct	taagaataac	atgtagaaca	42360
aagtacagta	ctgaatgatc	tttgttgaat	aaatatgaat	ggatattcaa	ataattaaaa	42420
atctcttaag	atctcccat	ctttacagga	tacagagaaa	actcgtaaat	atggcctgac	42480
ttttaccttt	gcagccttat	ccaaactctg	tgggtcaagac	aaacagggtg	tccttatact	42540
tacaacgtcc	ccctttgcct	acaaagctct	tctcatgact	ctttgcctat	cttaagttca	42600
cctatctgtc	aaatctctgg	gaatgcaaca	tttctcaag	gtagccttct	ctcctcccaa	42660
actagaacaa	attcttccctg	gggcattagg	tttttattgc	actgtatgtc	tcttcttcac	42720
agcaatcaca	gttccaatgt	tatatgtgta	ttcttagttg	atltgtttct	ttccaccttt	42780
agactataac	cttctaaggg	gtcacacata	atatcgatca	tcagttgtat	cccttgtgca	42840
tagcacaggg	catggcaggc	aaatatgtgt	gtaaataaac	ttgttgaatg	aatcaatgag	42900
acacactttt	cttaccocaa	gtataatggc	aggataacat	ttatcaatct	attgtctctt	42960
gaaaaacaga	tatgatgtgc	ttaattttca	ttttacatct	caaataccaa	tgccctaagg	43020
attcagagtc	attttacaaa	tctttttgac	aaatgccttc	attaatcacc	acctgtttac	43080
aagtgtctaa	taacattttg	gttacattct	gtaacatttc	ctgcacttaa	tgtcatctct	43140
agaatactgg	ctaatatgaa	gcacctggac	ttcaggaaca	caaacctgaa	actaacacac	43200
caaactaaac	tgttatgtaa	atgacagaaa	tgacacattt	tggtctgcaa	catctctaga	43260
tggtcttttg	accaattcaa	cttttaaccac	taaaaaatcg	tcacctgact	atagtcattt	43320
tgagctcatg	ataaatgaat	tacagatgaa	aaataaatag	tttgatgaca	atcttttaca	43380
aagtttatct	tcaagaata	ccaccagtca	caggttattct	aggctcctat	caactttatt	43440
ggtcagggca	gacttcactt	ttcatgataa	ttatgttctg	aaaattctac	aaacttaatg	43500
attacaaaca	aaagtcatag	tttgctcata	aatcaggcct	aggctctggat	tctagttctt	43560
ccatttttca	tttgttcact	gaggcaagtg	acttaaaatt	ccctagcctc	agtttctctc	43620
catgtaaaat	cagataatga	ttcctattcc	taagatgggt	ttgaggcttc	aacaagataa	43680
gatgggcctc	actcaagcat	gctcagtact	ctgtctctct	ctctccgggt	atgcagaaa	43740
tctattagga	ttctgcaaag	taaaataaat	atttcagtaa	aaattatgcc	ctttattaat	43800
gaatctagat	tttcagattt	tccttaaaat	tacttagtaa	cttaagggtc	caaataattat	43860
agagatttgt	atctagtatt	ttaaagaaat	gaaagggtgt	aatcaaaatg	ctgcacaaat	43920
aaatgctaca	tttaacaaac	agaatatcac	aaccatacaa	actaatcaga	tataaagaag	43980
tcagcaacag	aaatctgatg	ttgccttttag	atcacacaat	taggcaaaca	aaaatagagt	44040
tccatcctcc	tttgggtcaag	gccatggttg	aagactgaat	accaaatagg	gaaataggaa	44100
aagccaggaa	atggcaaat	agcaaaaact	ggactcctta	atttttatat	tcattttcat	44160
atctcacttc	taaaacttta	attaaattca	ataaaaaacc	aaaatggaac	tgagataaag	44220
ccaaaaggaa	agttatgtag	gtcaaagtga	aacctatatt	gtccttaggc	tctttgttgc	44280
tttctgttta	aggaaaaact	gccaagtgc	cttgacacat	taaagatcaa	gcaggagggt	44340
ctgccgagag	tccccatctg	gcagccagggt	tttgtcaagc	aaatttttag	aattctctac	44400
cctccacttc	tctatctaat	tatagcactt	tataaaaaacc	attctctctc	tgtctctgtc	44460
tctctctctc	tctctctctc	acacacacac	acacacacac	acacacacac	acacacaccc	44520
tttctctctc	tctctctctg	aaacttatct	gtattataat	aacacaacac	taggtatgga	44580
ttaatctgac	aattttcccc	taaaacagaa	taaattcaaa	aaggaaaacc	tttctctgtg	44640
acacatgcac	tatattctga	caataataat	tcctaaatta	agtataatac	attttcccta	44700
caggagttta	aagaagttac	agtaaagaat	ctcttgtata	aatatatatg	ccagaacttg	44760
acccaaataa	gtgctgagag	gtataaatct	caaaacagtt	tccggactct	ttgtgaaatg	44820

72

```

tcttcagagt ctgcgatata ttttcttcaa cttaaattata caagtaagat attttgctgg 44880
gctgtgggaa tgccttacgg catgttactg tggagctcat ggtaaaatag aaagaatata 44940
aataattaaa ataaaaattga caaatgataa atgattttaat aaattagaaa ttcaaatgcc 45000
gggcactttt ctagaacctg gacacaaagc atgaacctaa caataacccc gccttcatga 45060
aaaatatgga ctatttgaaa attataacctg caacactaaa taaatattct tcattcttcc 45120
agtatatgga gatgtttact ttcaattaga caatttgctt tcctctctga acacatagtt 45180
atgtgatggc tctataaaaag attttaaaat aactatagaa ggaactattg gtaaagactg 45240
tgggatacta aaaatggcta caaagaaagt tatgacaaaa cctctgagtt tgaatggaag 45300
tcctactaga ttagagtcta agcctgtgac attatgcttc tggttcttgt tcttaaatgc 45360
ttttctcatt aatagtatgt aacttacttc ctggaatgcc attcattaaa aaaatattta 45420
atatttgcta aatgtcaata tttatgccag cactttttaa gtacagaaac atggagtttc 45480
tttacctcat gcaaatatgc tgtgagaaag acttaagagc ctattgccta ctttgtggta 45540
caacactgaa gactcaccat ccaaaacaaa cagacttagt aaattcttgt gatttgcagt 45600
agttctgttc tataagggtt ccacaaacac tgaaatcatc gctcctgggg gaatacaagg 45660
ttatgtttcc gtgagccctc ggtcacaaca tgttcattaa ctgatcaata cataaccttg 45720
ttctatgtgt gtttctgttt aaaaagagca cttcagtgct acatttggag tctgttttaa 45780
acagcaaaa cactaataaa aagcacaaaa atgtaaaaagc atggcactac atacactgtg 45840
acaagaaggg ttgtttatag tatgacagct gagacaagaa ggtagagcct cgctttgatc 45900
aacctctgct gggaaatgag catcaggtga atcaattttt caccactctg aatgaccgta 45960
aaagtgtccc aagtactgac tttgggttta cacataaatt ttagtaagca tgtgaatctg 46020
ccaatatgaa atctacaaat aatgagtacc aaatgcatat gagtcaaata tttcagtgcg 46080
gtatctgact tgattgccac tgaaagacac agtttggaaa acccctaata aataccgttt 46140
agttactatg cagacaaaaga gttctacact agagtgcctc aattaagatg tctgaggctt 46200
tcataaatgg atgtttttta aaatgttatt tcctacctga tatattctaa aggggatata 46260
acgaaatcca ttttcttctg caggatattc catgagtttc cgattgatgg cccaaaactg 46320
gtcaaatctg tctgtaatat 46340

```

<210> 67
 <211> 773
 <212> ADN
 <213> Homo sapiens

```

<400> 67
actgagagac aggactagct ggatttccta ggctgactaa gaatccctaa gcctagctgg 60
gaagggtgacc acatccacct ttaaaccacgg ggcttgcaac ttagctcaca cctgaccaag 120
gaagggtgacc acaccctcct ttaaaccacag agcttgtaac tcagctcaca cccgaccaat 180
caggtagtaa agagagctca ctaaaatacc aattaggcta aaaacaggag gtaaagaaat 240
aatcaaatca tctatcgctt gagagcacag ggggaggggac aatgatcggg atataaaccc 300
aggcatttga gccagatcag gtaaccctct ttgggtcccc tcacactgta tgggagctct 360
gttttcactc tattaatctt tgcaactgca cactcttctg gtccatgttt gttccggctc 420
aagctgagct tttgctcgcc gtccaccact gctgaatgcc gccattgcag acctgccctt 480
gacttccacc cctccggatc cggcagagtg tccgctgcac tcctgatcca gcgaggcacc 540
cattgccact cccgatcagg ctaaaggctt gccattgttc ctgcacagct aagtgcctgg 600
gttcatccta atcaggctga aactggctcg ctgggttcca cggttctctt ccatgactca 660
cagcttctaa tagagctata aactcacca catggcccaa ggttccattc gttggaatcc 720
atgaggccaa gaaccccagg tcagagaata aaaggcccgcc ccatcttgg gag 773

```

<210> 68
 <211> 10
 <212> PRT
 <213> Homo sapiens

```

<400> 68
Phe Leu Gly Glu Glu Cys Cys Tyr Tyr Val
  1                               5               10

```

73

<210> 69
<211> 10
<212> PRT
<213> Homo sapiens

<400> 69
Leu Leu Phe Gly Pro Cys Ile Phe Asn Leu
1 5 10

<210> 70
<211> 10
<212> PRT
<213> Homo sapiens

<400> 70
Cys Leu Pro Leu Asn Phe Arg Pro Tyr Val
1 5 10

<210> 71
<211> 10
<212> PRT
<213> Homo sapiens

<400> 71
Gly Leu Leu Ser Gln Trp Met Pro Trp Ile
1 5 10

<210> 72
<211> 9
<212> PRT
<213> Homo sapiens

<400> 72
Cys Leu Pro Ser Gly Ile Phe Phe Val
1 5

<210> 73
<211> 9
<212> PRT
<213> Homo sapiens

<400> 73
Trp Met Pro Trp Ile Leu Pro Phe Leu
1 5

74

<210> 74
<211> 10
<212> PRT
<213> Homo sapiens

<400> 74
Ile Arg Trp Val Thr Pro Pro Thr Gln Ile
1 5 10

<210> 75
<211> 10
<212> PRT
<213> Homo sapiens

<400> 75
Leu Arg Asn Thr Gly Pro Trp Gly Leu Leu
1 5 10

<210> 76
<211> 10
<212> PRT
<213> Homo sapiens

<400> 76
Leu Arg Thr His Thr Arg Leu Val Ser Leu
1 5 10

<210> 77
<211> 10
<212> PRT
<213> Homo sapiens

<400> 77
Lys Arg Val Pro Ile Leu Pro Phe Val Ile
1 5 10

<210> 78
<211> 10
<212> PRT
<213> Homo sapiens

<400> 78
Cys Arg Cys Met Thr Ser Ser Ser Pro Tyr
1 5 10

75

<210> 79
<211> 10
<212> PRT
<213> Homo sapiens

<400> 79
Thr Arg Val His Gly Thr Ser Ser Pro Tyr
1 5 10

<210> 80
<211> 10
<212> PRT
<213> Homo sapiens

<400> 80
Ala Arg Glu Lys His Val Lys Glu Val Ile
1 5 10

<210> 81
<211> 10
<212> PRT
<213> Homo sapiens

<400> 81
Ser Arg Ile Glu Ala Val Lys Leu Gln Met
1 5 10

<210> 82
<211> 10
<212> PRT
<213> Homo sapiens

<400> 82
Ser Gln Trp Met Pro Trp Ile Leu Pro Phe
1 5 10

<210> 83
<211> 9
<212> PRT
<213> Homo sapiens

<400> 83
Cys Tyr Tyr Val Asn Gln Ser Gly Ile
1 5

76

<210> 84
<211> 9
<212> PRT
<213> Homo sapiens

<400> 84
Phe Tyr Tyr Lys Leu Ser Gln Glu Leu
1 5

<210> 85
<211> 9
<212> PRT
<213> Homo sapiens

<400> 85
Thr Tyr Thr Thr Asn Ser Gln Cys Ile
1 5

<210> 86
<211> 9
<212> PRT
<213> Homo sapiens

<400> 86
Ser Phe Leu Val Pro Pro Met Thr Ile
1 5

<210> 87
<211> 9
<212> PRT
<213> Homo sapiens

<400> 87
Tyr Tyr Val Asn Gln Ser Gly Ile Val
1 5

<210> 88
<211> 9
<212> PRT
<213> Homo sapiens

<400> 88
Leu Phe Asn Thr Thr Leu Thr Gly Leu
1 5

77

<210> 89
<211> 9
<212> PRT
<213> Homo sapiens

<400> 89
Leu Phe Gly Pro Cys Ile Phe Asn Leu
1 5

<210> 90
<211> 9
<212> PRT
<213> Homo sapiens

<400> 90
Arg Trp Val Thr Pro Pro Thr Gln Ile
1 5

<210> 91
<211> 10
<212> PRT
<213> Homo sapiens

<400> 91
Leu Pro Phe Leu Gly Pro Leu Ala Ala Ile
1 5 10

<210> 92
<211> 10
<212> PRT
<213> Homo sapiens

<400> 92
Leu Pro Tyr His Ile Phe Leu Phe Thr Val
1 5 10

<210> 93
<211> 10
<212> PRT
<213> Homo sapiens

<400> 93
Gly Ala Leu Gly Thr Gly Ile Gly Gly Ile
1 5 10

78

<210> 94
<211> 10
<212> PRT
<213> Homo sapiens

<400> 94
Leu Pro Phe Val Ile Gly Ala Gly Val Leu
1 5 10

<210> 95
<211> 9
<212> PRT
<213> Homo sapiens

<400> 95
Arg Arg Pro Leu Asp Arg Pro Ala Ser
1 5

<210> 96
<211> 9
<212> PRT
<213> Homo sapiens

<400> 96
Phe Arg Pro Tyr Val Ser Ile Pro Val
1 5

<210> 97
<211> 9
<212> PRT
<213> Homo sapiens

<400> 97
Arg Arg Ala Leu Asp Leu Leu Thr Ala
1 5

<210> 98
<211> 9
<212> PRT
<213> Homo sapiens

<400> 98
Trp Arg Met Gln Arg Pro Gly Asn Ile
1 5

79

<210> 99
<211> 10
<212> PRT
<213> Homo sapiens

<400> 99
Asp Arg Ile Gln Arg Arg Ala Glu Glu Leu
1 5 10

<210> 100
<211> 10
<212> PRT
<213> Homo sapiens

<400> 100
Leu Arg Thr His Thr Arg Leu Val Ser Leu
1 5 10

<210> 101
<211> 10
<212> PRT
<213> Homo sapiens

<400> 101
Glu Arg Val Ala Asp Ser Leu Val Thr Leu
1 5 10

<210> 102
<211> 10
<212> PRT
<213> Homo sapiens

<400> 102
Leu Phe Gly Pro Cys Ile Phe Asn Leu Leu
1 5 10

<210> 103
<211> 10
<212> PRT
<213> Homo sapiens

<400> 103
Gln Phe Tyr Tyr Lys Leu Ser Gln Glu Leu
1 5 10

80

<210> 104
<211> 10
<212> PRT
<213> Homo sapiens

<400> 104
Gln Trp Met Pro Trp Ile Leu Pro Phe Leu
1 5 10

<210> 105
<211> 10
<212> PRT
<213> Homo sapiens

<400> 105
Cys Tyr Tyr Val Asn Gln Ser Gly Ile Val
1 5 10

<210> 106
<211> 10
<212> PRT
<213> Homo sapiens

<400> 106
Asn Phe Val Ser Ser Arg Ile Glu Ala Val
1 5 10

<210> 107
<211> 9
<212> PRT
<213> Homo sapiens

<400> 107
Gly Pro Leu Val Ser Asn Leu Glu Ile
1 5

<210> 108
<211> 9
<212> PRT
<213> Homo sapiens

<400> 108
Leu Pro Leu Asn Phe Arg Pro Tyr Val
1 5

81

<210> 109
<211> 10
<212> PRT
<213> Homo sapiens

<400> 109
Leu Pro Phe Leu Gly Pro Leu Ala Ala Ile
1 5 10

<210> 110
<211> 10
<212> PRT
<213> Homo sapiens

<400> 110
Glu Pro Lys Met Gln Ser Lys Thr Lys Ile
1 5 10

<210> 111
<211> 10
<212> PRT
<213> Homo sapiens

<400> 111
Leu Pro Tyr His Ile Phe Leu Phe Thr Val
1 5 10

<210> 112
<211> 9
<212> PRT
<213> Homo sapiens

<400> 112
Arg Glu Lys His Val Lys Glu Val Ile
1 5

<210> 113
<211> 10
<212> PRT
<213> Homo sapiens

<400> 113
Lys Pro Arg Asn Lys Arg Val Pro Ile Leu
1 5 10

82

<210> 114
<211> 9
<212> PRT
<213> Homo sapiens

<400> 114
Val Val Leu Gln Asn Arg Arg Ala Leu
1 5

<210> 115
<211> 10
<212> PRT
<213> Homo sapiens

<400> 115
Ala Val Val Leu Gln Asn Arg Arg Ala Leu
1 5 10

<210> 116
<211> 9
<212> PRT
<213> Homo sapiens

<400> 116
Leu Pro Phe Val Ile Gly Ala Gly Val
1 5

<210> 117
<211> 9
<212> PRT
<213> Homo sapiens

<400> 117
Asp Leu Tyr Ser Tyr Val Ile Ser Lys
1 5

<210> 118
<211> 10
<212> PRT
<213> Homo sapiens

<400> 118
Thr Glu Gln Asp Leu Tyr Ser Tyr Val Ile
1 5 10

<210> 119
 <211> 2615
 <212> ADN
 <213> Homo sapiens

<400> 119
 gaattccggg aagccagacg gttaacacag acaaagtgtc gccgtgacac tcggccctcc 60
 agtgttgccg agaggcaaga gcagcgaccg cgcacctgtc cggccggagc tgggacgcgc 120
 gcccggggcg cgggacgaag cgaggaggga ccgccgaggc tgcccccaag tgtaactcca 180
 gcactgtgag gtttcaggga ttggcagagg ggaccaaggg gacatgaaaa tggacatgga 240
 ggatgccgat atgactctgt ggacagaggc tgagtttgaa gagaagtgtg catacattgt 300
 gaacgaccac ccctgggatt ctggtgtgta tggcgggtact tcggttcagg cggaggcatc 360
 cttaccaagg aatctgcttt tcaagtatgc caccaacagt gaagagggtta ttggagtgat 420
 gagtaaagaa tacataccaa agggcacacg ttttggaccc ctaatagggtg aaatctacac 480
 caatgacaca gttcctaaga acgccaacag gaaatatattt tggaggatct attccagagg 540
 ggagcttcac cacttcattg acggctttaa tgaagagaaa agcaactgga tgcgctatgt 600
 gaatcccgca cactctcccc gggagcaaaa cctggctgcy tgtcagaacg ggatgaacat 660
 ctacttctac accattaagc ccattccctgc caaccaggaa cttcttgtgt ggtattgtcg 720
 ggactttgca gaaaggcttc actaccctta tcccggagag ctgacaatga tgaatctcac 780
 acaaacacag agcagtctaa agcaaccgag cactgagaaa aatgaactct gcccaaagaa 840
 tgtcccaaag agagagtaca gcgtgaaaaga aatcctaaaa ttggactcca accctccaa 900
 aggaaaggac ctctaccgtt ctaacatttc accctcaca tcagaaaagg acctcgatga 960
 ctttagaaga cgtgggagcc ccgaaatgcc cttctaccct cgggtcgttt accccatccg 1020
 ggccctctg ccagaagact ttttgaaagc ttccctggcc tacgggatcg agagaccac 1080
 gtacatcact cgctcccca ttccatctc caccactcca agccctctg caagaagcag 1140
 ccccgacca agcctcaaga gctccagccc tcacagcagc cctgggaata cgggtgcccc 1200
 tgtggggccc ggctctcaag agcaccggga ctctacgct tacttgaacg cgtcctacgg 1260
 cacggaagg ttgggtcctt accctggcta cgcaccctg cccacctcc cgccagcttt 1320
 catccctcg tacaacgctc actacccaa gtctctcttg cccctctacg gcatgaattg 1380
 taatggcctg agcgtgtgta gcagcatgaa tggcatcaac aactttggcc tcttcccag 1440
 gctgtgccct gtctacagca atctcctcgg tgggggcagc ctgccccacc ccattgctca 1500
 cccacttct ctcccgagct cgctgccctc agatggagcc cggagggttg tccagccgga 1560
 gcatcccagg gaggtgcttg tcccggcgcc ccacagtgcc ttctccttta ccggggccgc 1620
 cgccagcatg aaggacaagg cctgtagccc cacaagcggg tctcccacgg cgggaacagc 1680
 cgccacggca gaacatgtgg tgcagcccaa agctacctca gcagcgatgg cagccccag 1740
 cagcgacgaa gccatgaatc tcattaaaaa caaaagaaac atgaccggct acaagaccct 1800
 tccctacccg ctgaagaagc agaaccggca gatcaagtac gaatgcaacg tttgcgccaa 1860
 gactttcggc cagctctcca atctgaaggt ccacctgaga gtgcacagtg gagaacggcc 1920
 tttcaaatgt cagacttgca acaagggtt tactcagctc gccacctgc agaaacacta 1980
 cctggtacac acgggagaaa agccacatga atgcccaggtc tgccacaaga gatttagcag 2040
 caccagcaat ctcaagaccc acctgcgact ccattctgga gagaaacat accaatgcaa 2100
 ggtgtgccct gccaaagttc ccagtttgt gcacctgaaa ctgcacaagc gtctgcacac 2160
 ccgggagcgg cccacaagt gctcccagtg ccacaagaac tacatccatc tctgtagcct 2220
 caaggttcac ctgaaaggga actgcgctgc ggccccggcg cctgggctgc ccttggaaaga 2280
 tctgacccga atcaatgaag aaatcgagaa gtttgacatc agtgacaatg ctgaccggct 2340
 cgaggacgtg gaggatgaca tcagtgtgat ctctgtagtg gagaaggaaa ttctggccgt 2400
 ggtcagaaaa gagaaagaag aaactggcct gaaagtgtct ttgcaaagaa acatggggaa 2460
 tggactctc tcctcagggt gcagccttta tgagtcatca gatctacccc tcatgaagtt 2520
 gcctcccagc aaccactac ctctggtacc tgtaaaggct aaacaagaaa cagttgaacc 2580
 aatggatcct taagattttc agaaaacact tattt 2615

<210> 120
 <211> 29
 <212> PRT
 <213> Homo sapiens

84

<400> 120

Leu Gln Asn Arg Arg Ala Leu Asp Leu Leu Thr Ala Glu Arg Gly Gly
1 5 10 15

Thr Cys Leu Phe Leu Gly Glu Glu Cys Cys Tyr Tyr Val
20 25

<210> 121

<211> 21

<212> ADN

<213> Homo sapiens

<400> 121

cttcaaacaa caaccaggag g

21

<210> 122

<211> 20

<212> ADN

<213> Homo sapiens

<400> 122

ttggggaggt tggccgacga

20

INTERNATIONAL SEARCH REPORT

International Application No

PCT/FR 99/01513

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/48 C12Q1/70 C07K14/15 A61K31/70

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C12Q C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 23755 A (BIO MERIEUX) 4 June 1998 (1998-06-04) Comparez nucléotides 1-1462 de SEQ ID NO:117 avec nucléotides 928-2390 de SEQ ID NO:1 de la présente demande; comparez SEQ ID NO:118 avec SEQ ID NOs:22 et 120 de la présente demande ---	1,3-12, 14-36
X	Database GenBank. Séquence HSAC 000064 Clone humain BAC RG083M05 de 7q21-7q22, séquence complet. 17 novembre 1996. XP002118730 Comparez nucléotides 28000-38500 d' AC00064 avec SEQ ID NO:3 --- -/--	1-4,13

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

18 October 1999

Date of mailing of the international search report

11.11.99

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Cupido, M

INTERNATIONAL SEARCH REPORT

International Application No

PCT/FR 99/01513

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ALLIEL PM ET AL: "Séquences rétrovirales endogènes anlogues à celle du nouveau rétrovirus MSRV associé à la sclérose en plaques (première partie)" COMPTES RENDUS DES SEANCES DE L'ACADEMIE DES SCIENCES SERIE III: SCIENCES DE LA VIE., vol. 321, no. 6, June 1998 (1998-06), pages 495-499, XP002101380 MONTREUIL FR figures 2,3 ---	1,3-12, 14-36
X	Database GenBank Séquence AC X93499 mRNA de H. sapiens pour la protéine rab7 10 février 1997 XP002119234 & VITELLI R ET AL: "Molecular cloning and expression analysis of the human rab7 GTP-ase complementary deoxyribonucleic acid" BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 229, no. 3, 1996, pages 887-890, ORLANDO, FL US ---	1-4
X	FR 2 737 500 A (BIO MERIEUX) 7 February 1997 (1997-02-07) cited in the application the whole document ---	1,3-12, 14-36
P,X	WO 99 02666 A (BIO MERIEUX) 21 January 1999 (1999-01-21) Comparez SEQ ID NOs 130, 117, 114 et 120 avec SEQ ID NOs: 1-3 de cette demande et SEQ ID NO:118 avec SEQ ID NOs:22 et 120 de cette demande ---	1,3-12, 14-36
P,X	WO 99 02696 A (BIO MERIEUX ;BESEME FREDERIC (FR); BLOND JEAN LUC (FR); BOUTON OLI) 21 January 1999 (1999-01-21) Comparez SEQ ID NOs: 4, 5, 7, 9 et 11 avec SEQ ID NOs: 1-3 de cette demande. ---	1,3-12, 14-36
P,X	ALLIEL PM ET AL: "Rétrovirus endogènes et sclérose en plaques. II. HERV-7q" COMPTES RENDUS DES SEANCES DE L'ACADEMIE DES SCIENCES SERIE III: SCIENCES DE LA VIE., vol. 321, no. 10, October 1998 (1998-10), pages 857-863, XP002101381 MONTREUIL FR the whole document --- -/--	1,3-12, 14-36

INTERNATIONAL SEARCH REPORT

International Application No

PCT/TR 99/01513

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 99 26972 A (GENETICS INSTITUTE, INC.) 3 June 1999 (1999-06-03) comparez SEQ ID NO:4 avec a.a. 131-668 de séquence 22 de cette demande ---	1-4
A	MITANI M ET AL: "Suppressive effect on polyclonal B-cell activation of a synthetic peptide homologous to a transmembrane component of oncogenic retrovirus" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 84, no. 1, January 1987 (1987-01), pages 237-240, XP002118729 WASHINGTON US cited in the application the whole document -----	24,25

INTERNATIONAL SEARCH REPORT

International application No.

PCT/FR 99/01513

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 8
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

See supplementary sheet INFORMATION FOLLOW-UP PCT/ISA/210

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See supplementary sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☒ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐
☐

The additional search fees were accompanied by the applicant's protest.

No protest accompanied the payment of additional search fees.

The International Searching Authority found several (groups of) inventions in the international application, namely:

1. Claims: 1, 9, 21-23, 26 (wholly), 28, 10-20 and 27-37 (partly)

Nucleic acid fragments derived from HERV-7q env, diagnostic reagents, diagnostic applications and kits, peptides, pharmaceutical compositions and applications, antibodies and corresponding transgenic animals.

2. Claims: 2-7, 10-20, 27-37 (all partly)

Nucleic acid fragments derived from HERV-7q gag, diagnostic reagents, diagnostic applications and kits, peptides, pharmaceutical compositions and applications, antibodies and corresponding transgenic animals.

3. Claims: 2-7, 10-20, 27-37 (all partly)

Human nucleic acid fragments similar to HERV-7q gag, diagnostic reagents, diagnostic applications and kits, peptides, pharmaceutical compositions and applications, antibodies and corresponding transgenic animals.

4. Claims: 24, 25

Compositions comprising a CKS-type motif, inasmuch as said compositions do not contain a sequence as per the first invention.

Continuation of Box I.2

Claim No: 8

Claim 8 concerns a very wide variety of compounds. A support basis as defined in PCT Article 6 and a description as defined in PCT Article 5 can however be found for only a very limited number of the claimed compounds. In the present case, the claims are so lacking in support basis and the disclosure of the invention in the description is so limited that it is impossible to carry out any significant search concerning the whole claimed spectrum.

The applicant's attention is drawn to the fact that claims concerning inventions in respect of which no search report has been established need not be the subject of a preliminary examination report (PCT Rule 66.1 (e)). The applicant is warned that the guideline adopted by the EPO acting in its capacity as International Preliminary Examining Authority is not to proceed with a preliminary examination of a subject matter unless a search has been carried out thereon. This position will remain unchanged, notwithstanding that the claims have or have not been modified, either after receiving the search report, or during any procedure under Chapter II.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/TR 99/01513

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9823755	A	04-06-1998	EP 0942987 A	22-09-1999
FR 2737500	A	07-02-1997	AU 6823296 A	05-03-1997
			BG 101355 A	30-12-1997
			BR 9606566 A	30-12-1997
			CA 2201282 A	20-02-1997
			CZ 9701357 A	17-06-1998
			EP 0789077 A	13-08-1997
			WO 9706260 A	20-02-1997
			HU 9900425 A	28-05-1999
			JP 11502416 T	02-03-1999
			NO 971493 A	03-06-1997
			NZ 316080 A	29-04-1999
			PL 319512 A	18-08-1997
			SK 56797 A	09-09-1998
WO 9902666	A	21-01-1999	FR 2765588 A	08-01-1999
			AU 8545098 A	08-02-1999
WO 9902696	A	21-01-1999	AU 8447098 A	08-02-1999
WO 9926972	A	03-06-1999	AU 1417899 A	15-06-1999

RAPPORT DE RECHERCHE INTERNATIONALE

Dem: Internationale No
PCT/FR 99/01513

A. CLASSEMENT DE L'OBJET DE LA DEMANDE

CIB 6 C12N15/48 C12Q1/70 C07K14/15 A61K31/70

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)

CIB 6 C12N C12Q C07K A61K

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés)

C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie °	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	WO 98 23755 A (BIO MERIEUX) 4 juin 1998 (1998-06-04) Comparez nucléotides 1-1462 de SEQ ID NO:117 avec nucléotides 928-2390 de SEQ ID NO:1 de la présente demande; comparez SEQ ID NO:118 avec SEQ ID NOs:22 et 120 de la présente demande ---	1,3-12, 14-36
X	Database GenBank. Séquence HSAC 000064 Clone humain BAC RG083M05 de 7q21-7q22, séquence complet. 17 novembre 1996. XP002118730 Comparez nucléotides 28000-38500 d' AC00064 avec SEQ ID NO:3 --- -/--	1-4,13



Voir la suite du cadre C pour la fin de la liste des documents



Les documents de familles de brevets sont indiqués en annexe

° Catégories spéciales de documents cités:

"A" document définissant l'état général de la technique, non
considéré comme particulièrement pertinent

"E" document antérieur, mais publié à la date de dépôt international
ou après cette date

"L" document pouvant jeter un doute sur une revendication de
priorité ou cité pour déterminer la date de publication d'une
autre citation ou pour une raison spéciale (telle qu'indiquée)

"O" document se référant à une divulgation orale, à un usage, à
une exposition ou tous autres moyens

"P" document publié avant la date de dépôt international, mais
postérieurement à la date de priorité revendiquée

"T" document ultérieur publié après la date de dépôt international ou la
date de priorité et n'appartenant pas à l'état de la
technique pertinent, mais cité pour comprendre le principe
ou la théorie constituant la base de l'invention

"X" document particulièrement pertinent; l'invention revendiquée ne peut
être considérée comme nouvelle ou comme impliquant une activité
inventive par rapport au document considéré isolément

"Y" document particulièrement pertinent; l'invention revendiquée
ne peut être considérée comme impliquant une activité inventive
lorsque le document est associé à un ou plusieurs autres
documents de même nature, cette combinaison étant évidente
pour une personne du métier

"&" document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

18 octobre 1999

Date d'expédition du présent rapport de recherche internationale

11.11.99

Nom et adresse postale de l'administration chargée de la recherche internationale

Office Européen des Brevets, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Fonctionnaire autorisé

Cupido, M

RAPPORT DE RECHERCHE INTERNATIONALE

Dem. Internationale No
PCT/FR 99/01513

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	<p>ALLIEL PM ET AL: "Séquences rétrovirales endogènes anlogues à celle du nouveau rétrovirus MSRV associé à la sclérose en plaques (première partie)" COMPTES RENDUS DES SEANCES DE L'ACADEMIE DES SCIENCES SERIE III: SCIENCES DE LA VIE., vol. 321, no. 6, juin 1998 (1998-06), pages 495-499, XP002101380 MONTREUIL FR figures 2,3</p> <p style="text-align: center;">---</p>	1,3-12, 14-36
X	<p>Database GenBank Séquence AC X93499 mRNA de H. sapiens pour la protéine rab7 10 février 1997 XP002119234 & VITELLI R ET AL: "Molecular cloning and expression analysis of the human rab7 GTP-ase complementary deoxyribonucleic acid" BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 229, no. 3, 1996, pages 887-890, ORLANDO, FL US</p> <p style="text-align: center;">---</p>	1-4
X	<p>FR 2 737 500 A (BIO MERIEUX) 7 février 1997 (1997-02-07) cité dans la demande le document en entier</p> <p style="text-align: center;">---</p>	1,3-12, 14-36
P,X	<p>WO 99 02666 A (BIO MERIEUX) 21 janvier 1999 (1999-01-21) Comparez SEQ ID NOs 130, 117, 114 et 120 avec SEQ ID NOs: 1-3 de cette demande et SEQ ID NO:118 avec SEQ ID NOs:22 et 120 de cette demande</p> <p style="text-align: center;">---</p>	1,3-12, 14-36
P,X	<p>WO 99 02696 A (BIO MERIEUX ;BESEME FREDERIC (FR); BLOND JEAN LUC (FR); BOUTON OLI) 21 janvier 1999 (1999-01-21) Comparez SEQ ID NOs: 4, 5, 7, 9 et 11 avec SEQ ID NOs: 1-3 de cette demande.</p> <p style="text-align: center;">---</p>	1,3-12, 14-36
P,X	<p>ALLIEL PM ET AL: "Rétrovirus endogènes et sclérose en plaques. II. HERV-7q" COMPTES RENDUS DES SEANCES DE L'ACADEMIE DES SCIENCES SERIE III: SCIENCES DE LA VIE., vol. 321, no. 10, octobre 1998 (1998-10), pages 857-863, XP002101381 MONTREUIL FR le document en entier</p> <p style="text-align: center;">---</p>	1,3-12, 14-36
	-/--	

RAPPORT DE RECHERCHE INTERNATIONALE

Dema Internationale No
PCT/FR 99/01513

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
P,X	WO 99 26972 A (GENETICS INSTITUTE, INC.) 3 juin 1999 (1999-06-03) comparez SEQ ID NO:4 avec a.a. 131-668 de séquence 22 de cette demande ---	1-4
A	MITANI M ET AL: "Suppressive effect on polyclonal B-cell activation of a synthetic peptide homologous to a transmembrane component of oncogenic retrovirus" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 84, no. 1, janvier 1987 (1987-01), pages 237-240, XP002118729 WASHINGTON US cité dans la demande le document en entier -----	24,25

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale n°
PCT/FR 99/01513

Cadre I Observations - lorsqu'il a été estimé que certaines revendications ne pouvaient pas faire l'objet d'une recherche (suite du point 1 de la première feuille)

Conformément à l'article 17.2)a), certaines revendications n'ont pas fait l'objet d'une recherche pour les motifs suivants:

1. ☐ Les revendications n^{os} se rapportent à un objet à l'égard duquel l'administration n'est pas tenue de procéder à la recherche, à savoir:

2. ☒ Les revendications n^{os} 8 se rapportent à des parties de la demande internationale qui ne remplissent pas suffisamment les conditions prescrites pour qu'une recherche significative puisse être effectuée, en particulier:
voir feuille supplémentaire SUITE DES RENSEIGNEMENTS PCT/ISA/210

3. ☐ Les revendications n^{os} sont des revendications dépendantes et ne sont pas rédigées conformément aux dispositions de la deuxième et de la troisième phrases de la règle 6.4.a).

Cadre II Observations - lorsqu'il y a absence d'unité de l'invention (suite du point 2 de la première feuille)

L'administration chargée de la recherche internationale a trouvé plusieurs inventions dans la demande internationale, à savoir:

voir feuille supplémentaire

1. ☐ Comme toutes les taxes additionnelles ont été payées dans les délais par le déposant, le présent rapport de recherche internationale porte sur toutes les revendications pouvant faire l'objet d'une recherche.

2. ☒ Comme toutes les recherches portant sur les revendications qui s'y prêtaient ont pu être effectuées sans effort particulier justifiant une taxe additionnelle, l'administration n'a sollicité le paiement d'aucune taxe de cette nature.

3. ☐ Comme une partie seulement des taxes additionnelles demandées a été payée dans les délais par le déposant, le présent rapport de recherche internationale ne porte que sur les revendications pour lesquelles les taxes ont été payées, à savoir les revendications n^{os}

4. ☐ Aucune taxe additionnelle demandée n'a été payée dans les délais par le déposant. En conséquence, le présent rapport de recherche internationale ne porte que sur l'invention mentionnée en premier lieu dans les revendications; elle est couverte par les revendications n^{os}

Remarque quant à la réserve

- ☐ Les taxes additionnelles étaient accompagnées d'une réserve de la part du déposant.
- ☐ Le paiement des taxes additionnelles n'était assorti d'aucune réserve.

SUITE DES RENSEIGNEMENTS INDIQUES SUR PCT/ISA/ 210

Suite du cadre I.2

Revendications nos.: 8

Le revendication 8 présente à trait à une très grande variété de composés. Un fondement au sens de L'Article 6 PCT et un exposé au sens de l'Article 5 PCT ne peut cependant être trouvé que pour un nombre très restreint de ces composés revendiqués. Dans le cas présent, les revendications manquent à un tel point de fondement et l'exposé de l'invention dans la description est si limité q'une recherche significative couvrant tout le spectre revendiqué est impossible.

L'attention du déposant est attirée sur le fait que les revendications, ou des parties de revendications, ayant trait aux inventions pour lesquelles aucun rapport de recherche n'a été établi ne peuvent faire obligatoirement l'objet d'un rapport préliminaire d'examen (Règle 66.1(e) PCT). Le déposant est averti que la ligne de conduite adoptée par l'OEB agissant en qualité d'administration chargée de l'examen préliminaire international est, normalement, de ne pas procéder à un examen préliminaire sur un sujet n'ayant pas fait l'objet d'une recherche. Cette attitude restera inchangée, indépendamment du fait que les revendications aient ou n'aient pas été modifiées, soit après la réception du rapport de recherche, soit pendant une quelconque procédure sous le Chapitre II.

SUITE DES RENSEIGNEMENTS INDIQUES SUR PCT/ISA/ 210

L'administration chargée de la recherche internationale a trouvé plusieurs (groupes d') inventions dans la demande internationale, à savoir:

1. revendications: 1, 9, 21-23, 26 (complet), 28, 10-20 et 27-37 (partiellement)

Fragments d'acide nucléique dérivé du HERV-7q env, réactifs de diagnostic, applications et kits diagnostiques, peptides, compositions et applications pharmaceutiques, anticorps et animaux transgéniques correspondants.

2. revendications: 2-7, 10-20, 27-37(tous partiellement)

Fragments d'acide nucléique dérivé du HERV-7q gag, réactifs de diagnostic, applications et kits diagnostiques, peptides, compositions et applications pharmaceutiques, anticorps et animaux transgéniques correspondants.

3. revendications: 2-7, 10-20, 27-37(tous partiellement)

Fragments d'acide nucléique humaines similaires à HERV-7q (SEQ ID NOs: 4-21 et 61), réactifs de diagnostic, applications et kits diagnostiques, peptides, compositions et applications pharmaceutiques, anticorps et animaux transgéniques correspondants.

4. revendications: 24, 25

Compositions comprenant un motif de type CKS, dans la mesure où ces compositions ne contiennent pas une séquence selon la première invention.

RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Demande internationale No

PCT/FR 99/01513

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO 9823755 A	04-06-1998	EP 0942987 A	22-09-1999
FR 2737500 A	07-02-1997	AU 6823296 A	05-03-1997
		BG 101355 A	30-12-1997
		BR 9606566 A	30-12-1997
		CA 2201282 A	20-02-1997
		CZ 9701357 A	17-06-1998
		EP 0789077 A	13-08-1997
		WO 9706260 A	20-02-1997
		HU 9900425 A	28-05-1999
		JP 11502416 T	02-03-1999
		NO 971493 A	03-06-1997
		NZ 316080 A	29-04-1999
		PL 319512 A	18-08-1997
		SK 56797 A	09-09-1998
WO 9902666 A	21-01-1999	FR 2765588 A	08-01-1999
		AU 8545098 A	08-02-1999
WO 9902696 A	21-01-1999	AU 8447098 A	08-02-1999
WO 9926972 A	03-06-1999	AU 1417899 A	15-06-1999